

SEARCH REQUEST FORM

Requestor's

Name: _____

Serial

Number: _____

Date: _____

Phone: _____

Art Unit: _____

Search Topic:

Please write a detailed statement of search topic. Describe specifically as possible the subject matter to be searched. Define any terms that may have a special meaning. Give examples or relevant citations, authors, keywords, etc., if known. For sequences, please attach a copy of the sequence. You may include a copy of the broadest and/or most relevant claim(s).

STAFF USE ONLY

Date completed: 08-29-02

Searcher: Beverly e-4994

Terminal time: 20

Elapsed time: _____

CPU time: _____

Total time: 25

Number of Searches: _____

Number of Databases: 1

Search Site

_____ STIC

_____ CM-1

_____ Pre-S

Type of Search

_____ N.A. Sequence

_____ A.A. Sequence

_____ Structure

_____ Bibliographic

Vendors

_____ IG

_____ STN

_____ Dialog

_____ APS

_____ Geninfo

_____ SDC

_____ DARC/Questel

☒ Other CGN

THIS PAGE BLANK (USPTO)

✓
STIC-Biotech/Ch mLib

From: Yu, Misook
Sent: Wednesday, August 21, 2002 8:24 AM
T : STIC-Biotech/ChemLib
Subject: 09/800,321

Please search SEQ ID NO:3 and 4. Thank you.

Examiner Misook Yu, Ph.D.
703-308-2454 (Phone)
Art Unit 1642
CM1-8E18 (Room)
CM1-8E12 (Mail Box)

Point of Contact:
Beverly Shears
Technical Info. Specialist
CM1 1E05 Tel: 308-4994

Searcher: _____
Phone: _____
Location: _____
Date Picked Up: _____
Date Completed: _____
Searcher Prep/Review: _____
Clerical: _____
Online time: _____

TYPE OF SEARCH:

NA Sequences: _____
AA Sequences: _____
Structures: _____
Bibliographic: _____
Litigation: _____
Full text: _____
Patent Family: _____
Other: _____

VENDOR/COST (where applic.)

STN: _____
DIALOG: _____
Questel/Orbit: _____
DRLink: _____
Lexis/Nexis: _____
Sequence Sys.: _____
WWW/Internet: _____
Other (specify): _____

THIS PAGE BLANK (USPTO)

XX WO200166746-A2.
PN 13-SEP-2001.
XX
XX
PF 05-MAR-2001: 2001WO-US07116.
XX
XX 03-MAR-2000: 2000US-0186606.
PR 06-MAR-2000: 2000US-0187247.
PR 06-MAR-2000: 2000US-0187248.
PR 06-MAR-2000: 2000US-0187249.
PR 06-MAR-2000: 2000US-0187250.
PR 06-MAR-2000: 2000US-0187253.
PR 06-MAR-2000: 2000US-0187295.
PR 06-MAR-2000: 2000US-0187296.
PR 07-MAR-2000: 2000US-0187563.
PR 21-JUL-2000: 2000US-0219854.
PR 24-JUL-2000: 2000US-0220263.
PR 31-JUL-2000: 2000US-0221942.
PR 21-DEC-2000: 2000US-0257600.
PR 08-JAN-2001: 2001US-0260285.
XX
XX (CURA-) CURAGEN CORP.
XX Padigaru M, Burgess CE, Majumder K, Mishra VS, Li L;
PI Baumgartner JC, Spytek KA, Tchernev VT;
XX N-PSDB: AAD17016.
DR MPI: 2001-565586/53.
XX
XX Novel polypeptides designated as NOVX polypeptides, useful in
PT detection, prevention and treatment of e.g. Parkinson's disease and
PT Cancer -
XX
XX
PS Claim 1: Page 10; 174pp; English.
XX
XX The invention relates to novel G-protein coupled-receptor (GPCR) related
CC polypeptides and their corresponding nucleotides, referred as NOVX or
CC NOV1, NOV2, NOV3, NOV4, NOV5, NOV6, NOV7, NOV8, NOV9 and NOV10. NOVX and
CC its Ab are useful for treating or preventing a NOVX-associated disorder,
CC such as disorder related to cell signal processing and metabolic signal
CC pathway modulation, e.g. diabetes. Pharmaceutical composition comprising
CC NOVX sequence is used to treat or to prevent disorders or syndromes
CC including metabolic disturbances associated with obesity, anorexia,
CC wasting disorders associated with chronic diseases, infectious diseases
CC (particularly infections caused by HIV-1 or HIV-2), cancer-associated
CC cachexia, cancer, Parkinson's disease, Alzheimer's disease, immune
CC disorders, neurodegenerative diseases and haematopoietic disorders. NOVX
CC sequence is also useful for treating developmental diseases, major
CC histocompatibility complex (MHC) II and III diseases (immune diseases),
CC noninsulin-dependent diabetes mellitus (NIDDM1), bulimia, hypotension,
CC acute heart failure, multiple sclerosis, hypertension, osteoporosis,
CC Crohn's disease, Albritght Hereditary Osteodystrophy, angina pectoris,
CC myocardial infarction, asthma, allergies, benign prostatic hypertrophy
CC and psychotic and neurological disorders, including anxiety, delirium,
CC schizophrenia, manic depression, dementia, severe mental retardation,
CC dentatorubralpallidolysian atrophy (DRPLA) hypophosphataemic rickets,
CC autosomal dominant (2) acrocallosal syndrome and dyskinesias, such as
CC Huntington's disease or Gilles de la Tourette syndrome and/or other
CC pathologies. NOVX DNA is used in mapping the chromosomal location of
CC NOVX gene and in forensic biology. The present sequence is GPCR NOV1b
CC protein.
CC
XX Sequence 313 AA:

Query Match 100.0%; Score 1607; DB 22: Length 313;
Best Local Similarity 100.0%; Pred. No. 4.8e-170;
Matches 313; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNWVNSIIIOEFLGPSRPMLEPLVAVLEISVTYIRGNLTIIVSLDRKLRPMW 60
DB 1 mnwvnsiiqeflllgfsdtpwlefpilvvlflsvtlvtlgnllilvslrdcklnpmv 60

QY 61 FFLTNLSLDLCYTTCTVPQMLVNLCSIRKVISYRGCAQLFTLALGATEVLLAVMSF 120
DB 61 ffltnslldlcyttctvpqmlvnlcsirkvisyrgcvaqlfflialgatevlllavmsf 120
QY 121 DMFVAICRPLHYSHVIMHQRCLDLAAASWYTGSSNSWMLSTLRLQPLCPPIYIDHFLCE 180
DB 121 dmfvaicrplhyshvimgqrclclaaaswvtgtsnswmlstlrlqplcdpyidhflce 180
QY 181 VPALKLSCVETTANEAELFLVSELFLILPTLLISYAFIVRAVLRIGSAGEGRKAFGT 240
DB 181 vpalklscvettanearaelflvseflfllpdlilslsyafivravlrigsagrkaftg 240
QY 241 CGSHLIVLSLEFYSTAVSVYLQPPSPSSKDGCKMVSLEFYGIAPMLNPLVTLRNKEVKEG 300
DB 241 cgsahlivslsfystavsvylqppspsskdqkmvslefylgiapmlnpllytlrnkevkeg 300
QY 301 FKRLVARVFLIKK 313
DB 301 fkrivarvflikk 313

RESULT 2
AAU04689
ID AAU04689 standard; Protein; 313 AA.
XX
AC AAU04689;
XX
DT 26-SEP-2001 (first entry)
XX
DE Novel human G-protein coupled receptor (NGPCR).
XX Human: novel G-protein coupled receptor; NGPCR; drug discovery;
KW diagnostic.
KM
XX Homo sapiens.
OS
XX WO200151634-A1.
PN
XX 19-JUL-2001.
PD
XX
PF 05-JAN-2001: 2001WO-US00589.
XX
XX 12-JAN-2000: 2000US-0175764.
PR
XX (LEXI-) LEXICON GENETICS INC.
PA
PI Walke DW, Wilganowski NL, Turner CA, Friedrich G, Abuin A;
PI Zambrowicz B, Sands AT;
XX N-PSDB: AAS08541.
DR MPI: 2001-442145/47.
XX
XX New human G-protein-coupled receptor and polynucleotides encoding the
PT receptor, useful in identifying, selecting or validating new molecular
PT targets for drug discovery and in diagnostic or prognostic assays -
XX
XX Claim 2: Page 61-62; 65pp; English.
XX
XX The sequence represents the amino acid sequence of novel human G-protein
CC coupled receptor (NGPCR). The NGPCR nucleic acid and protein are useful
CC in identifying, selecting or validating novel molecular targets for drug
CC discovery, as well as in diagnostic or prognostic assays. These are also
CC useful in microarrays or other assay formats, for screening collections
CC of genetic material from patients who have a particular medical condition
CC or for identifying mutations associated with a particular disease.
XX
XX Sequence 313 AA:

Query Match 99.1%; Score 1593; DB 22: Length 313;
Best Local Similarity 99.7%; Pred. No. 1.7e-168;
Matches 312; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MNWVNDSTIOEFTLLGSDRPWLEFPLLVLVFLISYVTYTGNTLTIIVSRDLTKLHPMY 60
 DB 1 mnwvndstioeftllgfsdrpwlefpvlvflisylvtfignltllivsrldtklhpmy 60
 QY 61 FFLTNLSLDLCYTCCTVCPQPMVLNLCIRKVISYRGCAOLFILAAGATEYLLAVMSF 120
 DB 61 ffltnlsldlcytcctvcpqpmvlncsrkvysrgcvaqlfilaagateyllavmsf 120
 QY 121 DWEVAICRPHYSYIMHQRCLQLAASWYTGFSNSWMLSTLLQLPLCDPYVIDHFLCE 180
 DB 121 dwevaicrphysyimhqrclqlaaswytgfsnswmlstllqlplcdpyvidhflce 180
 QY 181 VPALLKISCVEETANEAELFVSELFHLIPLTLLISYAFIVRAVLIRQSAEGRQKAFGT 240
 DB 181 vpallkiscvettaneaelfvseelfhlipltllisyaflvravlirgsaegrqkafgt 240
 QY 241 CGSHLIYVSLFYSYAVSVYIOPSPSSKDGKMWSLFYGIITAPMLNPLIYLRKKEVKEG 300
 DB 241 cgsqliyvslyfsyavsvyloppsskdgkmwslfygiitapmlnplilylrnkvekeg 300
 QY 301 FKRLVARVFLIKK 313
 DB 301 fkrivarvflikk 313

RESULT 3

AAAG71842 standard; Protein; 313 AA.

AAAG71842:

30-JUL-2001 (first entry)

Human olfactory receptor polypeptide, SEQ ID NO: 1523.

Human: olfactory receptor; OR; primary scent determination;

secondary scent determination; polypeptide library; odour receptor;

scent profile; scent fingerprint; scent representation.

Homo sapiens.

MO200127158-A2.

19-APR-2001.

06-OCT-2000; 2000WO-US27582.

08-OCT-1999; 99US-0158615.

24-FEB-2000; 2000US-0184809.

(DIGI-) DIGISCENTS.

(YEDA) YEDA RES & DEV CO LTD.

Bellenson J, Smith D, Lancet D, Glusman G, Fuchs T, Yanai I;

WPI; 2001-290713/30.

New polynucleotides which encode polypeptides involved in olfactory

sensation for identifying olfactory agonists and antagonists -

Claim 11; Page 972-973; 1857pp; English.

The present sequence is an olfactory receptor which is encoded by one of a number of novel polynucleotides. The polynucleotides can be used in screening for olfactory agonists and antagonists. The methods allow for the determination of primary scents and the identification of the odour receptors used to detect these primary scents. The methods also enable determination of secondary scents and the identification of combinations of odour receptors that are involved in detecting such secondary scents. This enables the construction of a scent representation also called a scent fingerprint or scent profile, which may be used to re-create and edit scents. Libraries of olfactory receptors are useful

for determining the interaction pattern of a composition with the CC receptors, and can be used for determining differences in the olfactory CC faculties of different individuals.

Sequence 313 AA:

Query Match 99.1%; Score 1593; DB 22; Length 313;
 Best Local Similarity 99.7%; Pred. No. 1,7e-168;
 Matches 312; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MNWVNDSTIOEFTLLGSDRPWLEFPLLVLVFLISYVTYTGNTLTIIVSRDLTKLHPMY 60
 DB 1 mnwvndstioeftllgfsdrpwlefpvlvflisylvtfignltllivsrldtklhpmy 60
 QY 61 FFLTNLSLDLCYTCCTVCPQPMVLNLCIRKVISYRGCAOLFILAAGATEYLLAVMSF 120
 DB 61 ffltnlsldlcytcctvcpqpmvlncsrkvysrgcvaqlfilaagateyllavmsf 120
 QY 121 DWEVAICRPHYSYIMHQRCLQLAASWYTGFSNSWMLSTLLQLPLCDPYVIDHFLCE 180
 DB 121 dwevaicrphysyimhqrclqlaaswytgfsnswmlstllqlplcdpyvidhflce 180
 QY 181 VPALLKISCVEETANEAELFVSELFHLIPLTLLISYAFIVRAVLIRQSAEGRQKAFGT 240
 DB 181 vpallkiscvettaneaelfvseelfhlipltllisyaflvravlirgsaegrqkafgt 240
 QY 241 CGSHLIYVSLFYSYAVSVYIOPSPSSKDGKMWSLFYGIITAPMLNPLIYLRKKEVKEG 300
 DB 241 cgsqliyvslyfsyavsvyloppsskdgkmwslfygiitapmlnplilylrnkvekeg 300
 QY 301 FKRLVARVFLIKK 313
 DB 301 fkrivarvflikk 313

RESULT 4

AAE09960 standard; Protein; 313 AA.

AAE09960:

29-NOV-2001 (first entry)

G-protein coupled-receptor (GPCR) NOVA1 protein.

G-protein coupled-receptor; GPCR; NOVA-associated disorder; obesity;

anorexia; wasting disorder; infection; cachexia; Parkinson's disease;

Alzheimer's disease; immune disorder; neurodegenerative disease; cancer;

anorectic; haematopoietic disorder; major histocompatibility complex;

MHCII; noninsulin-dependent diabetes mellitus; NIDDM; bulimia; asthma;

acute heart failure; hypotension; multiple sclerosis; hypertension;

osteoporosis; Crohn's disease; mental retardation; dementia; allergy;

angina pectoris; myocardial infarction; benign prostatic hypertrophy;

psychotic disorder; neurological disorder; anxiety; schizophrenia;

manic depression; delirium; Albright Hereditary Osteodystrophy; rickets;

dentatorubropallidolysian atrophy; DRPLA; haemostatic; anticonvulsant;

autosomal dominant (2) acrocallosal syndrome; dyskinesia; neuroleptic;

Huntington's disease; Gilles de la Tourette syndrome; neuroprotective;

neurotropic; antibacterial; protozoacide; fungicide; cytostatic; NOVA1;

vulnerable; cardiant; osteopathic; antitumoral; antidiabetic; virocidic;

human immunodeficiency virus; HIV; antidepressant; tranquilliser.

Unidentified.

Key location/Qualifiers

peptide 1..38 /label= signal_peptide

Protein 39..313 /note= "Mature GPCR NOVA1 protein"

MO200166746-A2.

PD		13-SEP-2001.	
XX			
PF		05-MAR-2001; 2001WO-US07116.	
PR		03-MAR-2000; 2000US-0186606.	
PR		06-MAR-2000; 2000US-0187247.	
PR		06-MAR-2000; 2000US-0187248.	
PR		06-MAR-2000; 2000US-0187249.	
PR		06-MAR-2000; 2000US-0187250.	
PR		06-MAR-2000; 2000US-0187253.	
PR		06-MAR-2000; 2000US-0187296.	
PR		07-MAR-2000; 2000US-0187563.	
PR		21-JUL-2000; 2000US-0219854.	
PR		24-JUL-2000; 2000US-0220263.	
PR		31-JUL-2000; 2000US-0221942.	
PR		21-DEC-2000; 2000US-0257600.	
PR		08-JAN-2001; 2001US-0260285.	
XX			
FA	(CURA-) CURAGEN CORP.		
P1	Padigar M, Burgess CE, Majumder K, Mishra VS, Li L;		
P1	Baumgartner JC, Spytek KA, Ichernev VI;		
DR	WPI: 2001-565586/63.		
DR	N-PSDB; AAD17015.		
PT	Novel polypeptides designated as NOXY polypeptides, useful in		
PT	detection, prevention and treatment of e.g. Parkinson's disease and		
PT	Cancer -		
XX			
PS	Claim 1; Page 9; 174pp; English.		
XX			
CC	The invention relates to novel G-protein coupled-receptor (GPCR) related		
CC	polypeptides and their corresponding nucleotides, referred as NOXY or		
CC	NOV1, NOV2, NOV3, NOV4, NOV5, NOV6, NOV7, NOV8, NOV9 and NOV10. NOXY and		
CC	its Ab are useful for treating or preventing a NOXY-associated disorder,		
CC	such as disorder related to cell signal processing and metabolic signal		
CC	pathway modulation, e.g. diabetes. Pharmaceutical composition comprising		
CC	NOXY sequence is used to treat or to prevent disorders or syndromes		
CC	including metabolic disturbances associated with obesity, anorexia,		
CC	wasting disorders associated with chronic diseases, infectious diseases		
CC	(particularly infections caused by HIV-1 or HIV-2), cancer-associated		
CC	cachexia, cancer, Parkinson's disease, Alzheimer's disease, immune		
CC	disorders, neurodegenerative diseases and haematopoietic disorders. NOXY		
CC	sequence is also useful for treating developmental diseases, major		
CC	histocompatibility complex (MHC) II and III diseases (immune diseases),		
CC	nonsulin-dependent diabetes mellitus (NIDDM), bulimia, hypotension,		
CC	acute heart failure, multiple sclerosis, hypertension, osteoporosis,		
CC	Crohn's disease, Albritcht Hereditary Osteodystrophy, angina pectoris,		
CC	myocardial infarction, asthma, allergies, benign prostatic hypertrophy		
CC	and psychotic and neurological disorders, including anxiety, delirium,		
CC	schizophrenia, manic depression, dementia, severe mental retardation,		
CC	dentatorubropallidoluysian atrophy (DRPLA) hypophosphataemic rickets,		
CC	autosomal dominant (2) acrocallosal syndrome and dyskinesias, such as		
CC	Huntington's disease or Gilles de la Tourette syndrome and/or other		
CC	pathologies. NOXY DNA is used in mapping the chromosomal location of		
CC	NOXY gene and in forensic biology. The present sequence is GPCR NOV1a		
XX			
XX			
SO	Sequence 313 AA;		
QY	Query Match 99.1%; Score 1593; DB 22; Length 313;		
Dd	Best Local Similarity 99.7%; Pred. No. 1.7e+166;		
Matches	312; Conservative 0; Mismatches 1; Indels 0; Gaps 0;		
QY	1 MNMVNDIIQIEFLLGFSDFRPWLLEPLLVEFLISYVTIIFGNLITILVSRLDTKLTHTPMY 60		
Dd	1 mnmvndiiqieflllgfdrplwlepllveflisytvtifgnltlilvsrltdklthtpmf 60		
QY	FFLRNLSLDLCTTTCYTPQMVLVNCSTRKYSTINGCYAAOLFIFLAGATEYELLVAWSF 120		
Dd	fflrnlslldcyytctcytpqmvlvncstrkystingcyaaolfiflagateyellvaWSF 120		

[illegible]

```

PR 08-JAN-2001: 2001us-0260285.
PA (CURA-) CURAGEN CORP.
XX
XX Padigar M, Burgess CE, Majumder K, Mishra VS, Li L;
PI Baumgartner JC, Spytek KA, Tchernev VT;
XX
XX WPI: 2001-565586/63.
DR N-PSDB; AAD17017.
XX
XX Novel polypeptides designated as NOVX polypeptides, useful in
PT detection, prevention and treatment of e.g. Parkinson's disease and
PT Cancer -
XX
XX Claim 1; Page 12; 174pp; English.
XX
XX The invention relates to novel G-protein coupled-receptor (GPCR) related
CC polypeptides and their corresponding nucleotides, referred as NOVX or
CC NOV1, NOV2, NOV3, NOV4, NOV5, NOV6, NOV7, NOV8, NOV9 and NOV10. NOVX and
CC its Ab are useful for treating or preventing a NOVX-associated disorder,
CC such as disorder related to cell signal processing and metabolic signal
CC pathway modulation, e.g. diabetes. Pharmaceutical composition comprising
CC NOVX sequence is used to treat or to prevent disorders or syndromes
CC including metabolic disturbances associated with obesity, anorexia,
CC wasting disorders associated with chronic diseases, infectious diseases
CC (particularly infections caused by HIV-1 or HIV-2), cancer-associated
CC cachexia, cancer, Parkinson's disease, Alzheimer's disease, immune
CC disorders, neurodegenerative diseases and haematopoietic disorders. NOVX
CC sequence is also useful for treating developmental diseases, major
CC histocompatibility complex (MHC) II and III diseases (immune diseases),
CC noninsulin-dependent diabetes mellitus (NIDDM1), bulimia, hypotension,
CC acute heart failure, multiple sclerosis, hypertension, osteoporosis,
CC Crohn's disease, Albright Hereditary Osteodystrophy, angina pectoris,
CC myocardial infarction, asthma, allergies, benign prostatic hypertrophy
CC and psychotic and neurological disorders, including anxiety, delirium,
CC schizophrenia, manic depression, dementia, severe mental retardation,
CC dementorubopallidolysian atrophy (DRPLA) hypophosphataemic rickets,
CC autosomal dominant (?) acrocallosal syndrome and dyskinesias, such as
CC Huntington's disease or Gilles de la Tourette syndrome and/or other
CC pathologies. NOVX DNA is used in mapping the chromosomal location of
CC CC gene and in forensic biology. The present sequence is GPCR NOV1C
XX
XX
XX Sequence 313 AA:
SQ

```

```

Query Match          98.7%; Score 1586; DB 22; Length 313;
Best Local Similarity 99.0%; Pred. No. 1e-167;
Matches 310; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 1 MNWVNDIIIOEFILGFSDBRPWFEPPLVFLISYTYTIFGNLTIIIVSRDLTKLHPMY 60
DB 1 mnnwnndiioefilgfsdbrpwlfepllvflisyytytignltiivsrldtklhpmymy 60
QY 61 FFLTNLSLDLCYTTCTVPMQNLVNLCSIRKVISYRGCAOLFPLAGATGYLLAVMSF 120
DB 61 flfntnlsldlcyttctvpmqlvnlcsirkvisyrgcvaolfplagatgyllavmsf 120
QY 121 DMFVAICRPLHYIVIMQRCICLQLAASWYTGFSNYSWLTSLTLQPLCPDYVIDHFLCE 180
DB 121 dmfvaircrlphysivimqrcicqlaaswytgfsnyswltstltqlpdcpyvidhflice 180
QY 181 VPALKISCVETTANAELELVSEFLHLPVLTLLISYAFIVRAVLRIGSAEGRQAKFGT 240
DB 181 vpalakisvettanaelelvseflhlpvltllisyaativravlrigsaegrqakfgt 240
QY 241 CGSHLIYVSLEFYTAASVYIOPSPSSKDGKMWLSLFGIAPMLNPLITLRKKEVKEG 300
DB 241 cgshtliyvslefytavsvyiqppssskdrgkmvslfygiaplmlplitylrmrevkeg 300
QY 301 FKRLVARVFLIKK 313
DB 301 fkrivarvflikk 313

```

```

RESULT 6
ID AAG72915
XX AAG72915 standard; Protein; 314 AA.
XX
XX AAG72915;
XX
XX 30-JUL-2001 (first entry)
XX
XX Human olfactory receptor data exploratorium sequence, SEQ ID NO: 2597.
DE
XX
XX Human; olfactory receptor; OR; primary scent determination;
KW secondary scent determination; polypeptide library; odour receptor;
KW scent profile; scent fingerprint; scent representation;
KW human olfactory receptor data exploratorium; HORDE.
XX
XX Homo sapiens.
OS
XX
XX WO200127158-A2.
XX
XX 19-APR-2001.
XX
XX 06-OCT-2000: 2000MO-US27582.
XX
XX 08-OCT-1999: 99US-0158615.
XX
XX 24-FEB-2000: 2000US-0184809.
XX
XX (DIGI-) DIGISCENTS.
PA (YEDA ) YEDA RES & DEV CO LTD.
XX
XX Bellenson J, Smith D, Lancet D, Glusman G, Fuchs T, Yanai I;
PI WPI: 2001-290713/30.
XX
XX New polynucleotides which encode polypeptides involved in olfactory
PT sensation for identifying olfactory agonists and antagonists -
XX
XX Example 6; Page 1754-1755; 1857pp; English.
XX
XX The present sequence is a polypeptide from the human olfactory receptor
CC data exploratorium (HORDE). It was used as a query sequence in a database
CC search of olfactory receptor (OR)-like sequences. The invention relates
CC to isolated polynucleotides encoding polypeptides involved in olfactory
CC sensation. The polynucleotides can be used in screening for olfactory
CC agonists and antagonists. The methods allow for the determination of
CC primary scents and the identification of the odour receptors used to
CC detect these primary scents. The methods also enable determination of
CC secondary scents and the identification of combinations of odour
CC receptors that are involved in detecting such secondary scents. This
CC enables the construction of a scent representation (also called a scent
CC fingerprint or scent profile), which may be used to re-create and edit
CC scents. Libraries of olfactory receptors are useful for determining the
CC interaction pattern of a composition with the receptors, and can be
CC used for determining differences in the olfactory faculties of different
CC individuals.
XX
XX Sequence 314 AA:
SQ

```

```

Query Match          81.9%; Score 1316; DB 22; Length 314;
Best Local Similarity 81.8%; Pred. No. 1.e-137;
Matches 256; Conservative 25; Mismatches 32; Indels 0; Gaps 0;

```

```

QY 1 MNWVNDIIIOEFILGFSDBRPWFEPPLVFLISYTYTIFGNLTIIIVSRDLTKLHPMY 60
DB 1 mnnwnndiioefilgfsdbrpwlfepllvflisyytytignltiivsrldtklhpmymy 60
QY 2 mnnwnkvsyvgqflllvsdbrpwlfepllvflisyytytignltiivsnvdfllhpmymy 61
DB 2 mnnwnkvsyvgqflllvsdbrpwlfepllvflisyytytignltiivsnvdfllhpmymy 61
QY 61 FFLTNLSLDLCYTTCTVPMQNLVNLCSIRKVISYRGCAOLFPLAGATGYLLAVMSF 120
DB 61 flfntnlsldlcyttctvpmqlvnlcsirkvisyrgcvaolfplagatgyllavmsf 120
QY 62 FFLTNLSLDLCYTTCTVPMQNLVNLCSIRKVISYRGCAOLFPLAGATGYLLAVMSF 121
DB 62 flfntnlsldlcyttctvpmqlvnlcsirkvisyrgcvaolfplagatgyllavmsf 121
QY 121 DMFVAICRPLHYIVIMQRCICLQLAASWYTGFSNYSWLTSLTLQPLCPDYVIDHFLCE 180
DB 121 dmfvaircrlphysivimqrcicqlaaswytgfsnyswltstltqlpdcpyvidhflice 180

```

```

Db      122 drfvaicrphylsilmqricfglaaswsgfnsvglstwtlkmplcgkhkevdfhfc 181
QY      181 VPAALKKSCVETTANAEALFLVSELPFLIPLTLLISYAFIVRAVLRIOGAEGROKAFGT 240
Db      182 vpallkiscvdtaneaeelffivfillipvtlllisyafivgavirigaegqrfgt 241
QY      241 CGSHLIYVSLFYSTAVSVYLQPPSPSSKDGKMYSLFYGIAPMLNPLIYTLRNKEVEKG 300
Db      242 cgsahlivslfygtalsmylqppspsskdrkmvslfcgllapmlnpllytlrnkvekea 301
QY      301 FKRUVARVFLIKK 313
Db      302 fkrllqvflfikk 314

RESULT 7
AAG71493
ID      AAG71493 standard; Protein; 357 AA.
XX
AC      AAG71493;
XX
DT      30-JUL-2001 (first entry)
XX
DE      Human olfactory receptor polypeptide, SEQ ID NO: 1174.
XX
KM      Human: olfactory receptor; OR: primary scent determination;
KW      secondary scent determination; polypeptide library; odour receptor;
XX      scent profile; scent fingerprint; scent representation.
XX
OS      Homo sapiens.
XX
PN      WO200127158-A2.
XX
PD      19-APR-2001.
XX
PE      06-OCT-2000; 2000MO-US27582.
XX
PR      08-OCT-1999; 99US-0158615.
XX
PR      24-FEB-2000; 2000US-0184809.
XX
PA      (DIGI-) DIGISCENTS.
XX      (YEDA ) YEDA RES & DEV CO LTD.
XX
PI      Bellenson J, Smith D, Lancet D, Glusman G, Fuchs T, Yanai I;
XX
DR      WPI; 2001-290713/30.
XX
PT      New polynucleotides which encode polypeptides involved in olfactory
XX      sensation for identifying olfactory agonists and antagonists -
XX
PS      Claim 11; Page 690-691; 1857pp; English.
XX
CC      The present sequence is an olfactory receptor which is encoded by
CC      one of a number of novel polynucleotides. The polynucleotides can be
CC      used in screening for olfactory agonists and antagonists. The methods
CC      allow for the determination of primary scents and the identification
CC      of the odour receptors used to detect these primary scents. The methods
CC      also enable determination of secondary scents and the identification of
CC      combinations of odour receptors that are involved in detecting such
CC      secondary scents. This enables the construction of a scent representation
CC      (also called a scent fingerprint or scent profile), which may be used to
CC      re-create and edit scents. Libraries of olfactory receptors are useful
CC      for determining the interaction pattern of a composition with the
CC      receptors, and can be used for determining differences in the olfactory
CC      faculties of different individuals.
XX
SQ      Sequence 357 AA;

```

Query Match 80.6%; Score 1295; DB 22; Length 357;
 Best Local Similarity 81.3%; Pred. No. 2.7e-135;
 Matches 252; Conservative 25; Mismatches 33; Indels 0; Gaps 0;

```

QY      1 MNWVNDIIIOEFLIGSDRPWLEFPLLVFLISYTVIFGNLTIIIVSRUDTFLHTPMY 60
Db      1 mnnwnkxvpgfeilllvdsdpwleiprfmflfisyllitfgnltlllvshvdfllhpm 60
QY      61 FFLTNLSLDLCYTTCYTPQMLVNLCSIRKVISYRGCAVQDLFFLALGATEYLLAVMSF 120
Db      61 ffltnslldlcytstevpmlvnlcntrkvisyrgvcaqlfflaigsteclllavmcf 120
QY      121 DMFAIRCPHLYSIMORCLQLAASWYTGFSNSWLSLTLTQLDLPPIVDHFLCE 180
Db      121 drfvaicrphylsilmqricfglaaswsgfnsvglstwtlkmplcgkhkevdfhfc 180
QY      181 VPAALKKSCVETTANAEALFLVSELPFLIPLTLLISYAFIVRAVLRIOGAEGROKAFGT 240
Db      181 vpallkiscvdtaneaeelffivfillipvtlllisyafivgavirigaegqrfgt 240
QY      241 CGSHLIYVSLFYSTAVSVYLQPPSPSSKDGKMYSLFYGIAPMLNPLIYTLRNKEVEKG 300
Db      241 cgsahlivslfygtalsmylqppspsskdrkmvslfcgllapmlnpllytlrnkvekea 300
QY      301 FKRUVARVFL 310
Db      301 fkrllqvksll 310

RESULT 8
AAG72074
ID      AAG72074 standard; Protein; 357 AA.
XX
AC      AAG72074;
XX
DT      30-JUL-2001 (first entry)
XX
DE      Human olfactory receptor polypeptide, SEQ ID NO: 1755.
XX
KM      Human: olfactory receptor; OR: primary scent determination;
KW      secondary scent determination; polypeptide library; odour receptor;
XX      scent profile; scent fingerprint; scent representation.
XX
OS      Homo sapiens.
XX
PN      WO200127158-A2.
XX
PD      19-APR-2001.
XX
PE      06-OCT-2000; 2000MO-US27582.
XX
PR      08-OCT-1999; 99US-0158615.
XX
PR      24-FEB-2000; 2000US-0184809.
XX
PA      (DIGI-) DIGISCENTS.
XX      (YEDA ) YEDA RES & DEV CO LTD.
XX
PI      Bellenson J, Smith D, Lancet D, Glusman G, Fuchs T, Yanai I;
XX
DR      WPI; 2001-290713/30.
XX
PT      New polynucleotides which encode polypeptides involved in olfactory
XX      sensation for identifying olfactory agonists and antagonists -
XX
PS      Claim 11; Page 1150-1151; 1857pp; English.
XX
CC      The present sequence is an olfactory receptor which is encoded by
CC      one of a number of novel polynucleotides. The polynucleotides can be
CC      used in screening for olfactory agonists and antagonists. The methods
CC      allow for the determination of primary scents and the identification
CC      of the odour receptors used to detect these primary scents. The methods
CC      also enable determination of secondary scents and the identification of
CC      combinations of odour receptors that are involved in detecting such
CC      secondary scents. This enables the construction of a scent representation
CC      (also called a scent fingerprint or scent profile), which may be used to
CC      re-create and edit scents. Libraries of olfactory receptors are useful

```

SQ Sequence 357 AA;

Matches	252;	Conservative	25;	Mismatches	33;	Indels	0;	Gaps	0;
---------	------	--------------	-----	------------	-----	--------	----	------	----

61 FFLTNLSLLDCYTTCTVPQMLNLCIRKVISYRGVAQLFI¹²⁰
 QY FLAGATEYLLAVMSF

Db 121. drfvaicrplhysimhgrlcfqlaaswsgfsnsvlqstwtlkmplcghkevdhfice 180

QY 241 CGSHLIVSLFYSTAVSVYLQPPSPSSKDQGMVSLFYGIAPMLNPLIYTLRNKEVKEG 300

Db 301 fkrivaks11 310

XX
XX
2003.

KW scent profile; scent fingerprint; scent representation.

PN W0200127158-A2.

PD 19-APR-2001.

06-OCT-2000; 2000WO-US27582.

PR 08-OCT-1999; 99US-0158615.

XX
XX

PA (YEDA) YEDA RES & DEV CO LTD.
XX

PI Belenson J, Smith D, Lancet D, Glusman G, Fuchs T, Yanai I;
XY

DR WPI; 2001-290713/30.

P1 New polynucleotides which encode polypeptides involved in olfactory sensation for identification of factors associated with olfactory dysfunction

PS Example 6; Page 1815; 1857pp; English.

X

CC individuals.

Best Local Similarity 78.08; Pred. No. 2.3e-131;

[illegible]

.....

XX

XX

XX

XX

KW secondary scent determination; polypeptide library; odour receptor;

KW human olfactory receptor data exploratorium; HORDE.

OS Homo sapiens.

PN W0200127158-A2.

PD 19-APR-2001.

XX

PR

PR 24-FEB-2000; 2000US-0184809.
 XX (DIGI-) DIGISCENTS.
 PA (YEDA) YEDA RES & DEV CO LTD.
 XX
 PI Bellenson J, Smith D, Lancet D, Glusman G, Fuchs T, Yanai I;
 XX WPI; 2001-290713/30.
 DR
 XX New polynucleotides which encode polypeptides involved in olfactory
 PT sensation for identifying olfactory agonists and antagonists -
 XX
 PS Example 6; Page 1755; 1857pp; English.
 XX
 XX The present sequence is a polypeptide from the human olfactory receptor
 CC data exploratorium (HOREDO). It was used as a query sequence in a database
 CC search of olfactory receptor (OR)-like sequences. The invention relates
 CC to isolated polynucleotides encoding polypeptides involved in olfactory
 CC sensation. The polynucleotides can be used in screening for olfactory
 CC agonists and antagonists. The methods allow for the determination of
 CC primary scents and the identification of the odour receptors used to
 CC detect these primary scents. The methods also enable determination of
 CC secondary scents and the identification of combinations of odour
 CC receptors that are involved in detecting such secondary scents. This
 CC enables the construction of a scent representation (also called a scent
 CC fingerprint or scent profile), which may be used to re-create and edit
 CC scents. Libraries of olfactory receptors are useful for determining the
 CC interaction pattern of a composition with the receptors, and can be
 CC used for determining differences in the olfactory faculties of different
 CC individuals.
 CC
 XX Sequence 343 AA;
 SQ

Query Match 78.3%; Score 1259; DB 22; Length 343;
 Best Local Similarity 78.0%; Pred. No. 2.6e-131;
 Matches 244; Conservative 29; Mismatches 40; Indels 0; Gaps 0;

QY 1 MNWVNDIIOEFILGFSDRPMEFPLLYVFLISYVTIGENLIIIVSRDTRKLTMPY 60
 DB 5 msvenesistreflllgfsdprwleplfvvflvsgylltlfqnmlllvstidskrlhpepy 64
 QY 61 FELNLISLIDLCYTCCTVQPMVLNLCGRKYSRGCAQOIFLALGAPTEYLLAAMSF 120
 DB 65 flfnlnslldcyctstsvpqlmlncstrkvisyggcvqgflfllsgstecfllyvmsl 124
 QY 121 DMEVAICRPILHSYIMHQRCLOLAASWVTGFSNSWMLSTLTQLPCDPEYVIDHFCE 180
 DB 125 drfaicrphlhyvymqtrchhaacwlsfgfnsylvgtwltqmpcgykewdhffice 184
 QY 181 VPALKLSVETTANAEALFLVSELPHLIPFTLILISYAFIVRAVLRIOAEGROKAGT 240
 DB 185 vpalklscvdtaneaelffisvllfllpytlllisyafivgvlklrseccrkaftg 244
 QY 241 CGSLILVSVLFYTRAVSVYIQPPSSSKDQKMSLFYGIAPMLNPLIYLRKKEVKEG 300
 DB 245 cgslllvsvlflgcatlmylqppssskdrgkmslflgyltlpmlnplilylrneevkga 304
 QY 301 FKRLVAVRFLIRK 313
 DB 305 fkrllmkrllilgk 317

RESULT 11
 AAB43266
 ID AAB43266 standard; Protein; 313 AA.
 AC AAB43266;
 XX
 DT 08-FEB-2001 (first entry)
 XX
 DE Human ORFX ORF3030 polypeptide sequence SEQ ID NO: 6060.
 XX

KW Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;
 KW vulnery; antipsothic; antiparkinsonian; nootropic; neuroprotective;
 KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;
 KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
 KW hypotensive; dermatological; immunosuppressive; antinflammatory;
 KW antiviral; antibacterial; antifungal; antirheumatic; antithyroid;
 KW antianaemic; gene therapy; cancer; proliferative disorder; hypertension;
 KW neurodegenerative disorder; osteoarthritis; graft vs host disease;
 KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
 KW cholesterol ester storage; systemic lupus erythematosus; infection;
 KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
 KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
 KW bone damage; cartilage damage; antinflammatory disease; coagulation;
 KW thrombosis; contraceptive.
 KW
 XX Homo sapiens.
 OS
 XX
 PN W0200058473-A2.
 XX
 PD 05-OCT-2000.
 XX
 PF 31-MAR-2000; 2000MO-US08621.
 XX
 PR 31-MAR-1999; 99US-0127607.
 PR 02-APR-1999; 99US-0127636.
 PR 03-APR-1999; 99US-0127728.
 PR 30-MAR-2000; 2000US-0540763.
 XX
 PA (CURA-) CURAGEN CORP.
 XX
 PI Shinkets RA, Leach M;
 XX
 DR WPI; 2000-602362/57.
 DR N-PSDB; AAC77475.
 XX
 PT Novel nucleic acids and peptides derived from open reading frame X,
 PT useful for treating e.g. cancers, proliferative disorders,
 PT neurodegenerative disorders and cardiovascular disease -
 XX
 XX Claim 11; Page 5246-5247; 5507pp; English.
 PS
 CC AAC74446 to AAC7606 encode the proteins given in AAB40237 to AAB43397,
 CC which represent the human ORFX open reading frames 1 to 3161. The ORFX
 CC sequences have activities such as: cytostatic; hepatotropic; vulnery;
 CC antipsothic; antiparkinsonian; nootropic; neuroprotective;
 CC osteopathic; anticonvulsant; antiarthritic; immunosuppressant;
 CC immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;
 CC antidiabetic; hypotensive; dermatological; immunosuppressive;
 CC antinflammatory; antibacterial; antiviral; antifungal; antirheumatic;
 CC antithyroid; and antianaemic. The sequences can be used for determining
 CC the presence of or predisposition to, or preventing or treating
 CC pathological conditions associated with an ORFX-associated disorder. The
 CC nucleic acids can be used to express ORFX proteins in gene therapy
 CC vectors. The proteins and nucleic acids may be used to treat cancers,
 CC proliferative disorders, neurodegenerative disorders, osteoarthritis,
 CC graft vs host disease, cardiovascular disease, diabetes mellitus,
 CC hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
 CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
 CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,
 CC allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
 CC nocturnal haemoglobinuria, antinflammatory disease; to enhance
 CC coagulation; to inhibit thrombosis; and as a contraceptive.
 CC
 XX Sequence 313 AA;
 SQ

Query Match 70.3%; Score 1130; DB 21; Length 313;
 Best Local Similarity 70.3%; Pred. No. 5e-117;
 Matches 220; Conservative 34; Mismatches 59; Indels 0; Gaps 0;

QY 1 MNWVNDIIOEFILGFSDRPMEFPLLYVFLISYVTIGENLIIIVSRDTRKLTMPY 60
 DB 1 mnwvesspkflllgfsdprwleplfvvflvsgylltlfqnmlllvstidskrlhpepy 60

XX 08-OCT-1999; 9905-0158615.
PR 24-FEB-2000; 2000US-0184809.
XX
XX (DIGI-) DIGISCENTS.
PA (YEDA) YEDA RES & DEV CO LTD.
XX
XX Bellenson J, Smith D, Lancel D, Glusman G, Fuchs T, Yanai I;
XX WPI; 2001-290713/30.
XX
XX New polynucleotides which encode polypeptides involved in olfactory
PT sensation for identifying olfactory agonists and antagonists -
XX
XX
XX Claim 11; Page 1136-1137; 1857pp; English.
XX
XX The present sequence is an olfactory receptor which is encoded by
CC one of a number of novel polynucleotides. The polynucleotides can be
CC used in screening for olfactory agonists and antagonists. The methods
CC allow for the determination of primary scents and the identification
CC of the odour receptors used to detect these primary scents. The methods
CC also enable determination of secondary scents and the identification of
CC combinations of odour receptors that are involved in detecting such
CC secondary scents. This enables the construction of a scent representation
CC (also called a scent fingerprint or scent profile), which may be used to
CC re-create and edit scents. Libraries of olfactory receptors are useful
CC for determining the interaction pattern of a composition with the
CC receptors, and can be used for determining differences in the olfactory
CC faculties of different individuals.
XX
XX
SQ Sequence 216 AA;

Query Match 66.9%; Score 1075; DB 22; Length 216;
Best Local Similarity 98.6%; Pred. No. 4e-111;
Matches 213; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 68 LLDLCYTTCTVPQMLVNLCSIRKVIYIRGCVAGLFIPLALGATEYLLAVMSFDMFVAIC 127
DB 1 lldlcyttctvpqmlvnlcsirkviyirgcvaqllfiflaigateyllavmsfdrfvaic 60
QY 128 RPLHYSVIMHORLCLQIAAASWMTGFSNSVWLSTLTQLPLCDPYVIDHFIQCEVPALIKL 187
DB 61 rplhysvimhgrlclqlaasrvtgfsnswlsltlqlplcdpyvidhfiqcevpallkl 120
QY 188 SCVETTANAEELFVSELFHPLPTLLISYAFIVRAVLRIOAEGROKAFGCGSHLIV 247
DB 121 scvettaneaelflvselfhplptlllisyafivravlrirgseagrqkafgctgshliv 180
QY 248 VSLFYSTAVSYRLQPPSPSSKDGKMSLFYGIAP 283
DB 181 vslfnstavsylvlqppspsskdgkmslfygiiap 216

Search completed: August 25, 2002, 20:07:13
Job time: 4206 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 25, 2002, 18:57:37 ; Search time 32.69 Seconds

(without alignments)
233.870 Million cell updates/sec

Title: US-09-800-321a-4

Perfect score: 1607
Sequence: 1 MMWVDSIIQEFILIGFSDR.....NKEVGEFRRLVAVFLINK 313

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 231628 segs, 2442594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents .AA:*
1: /cgn2.6/prodata/2/1aa/5a_COMB.pep.*
2: /cgn2.6/prodata/2/1aa/5b_COMB.pep.*
3: /cgn2.6/prodata/2/1aa/6a_COMB.pep.*
4: /cgn2.6/prodata/2/1aa/6b_COMB.pep.*
5: /cgn2.6/prodata/2/1aa/PCTUS_COMB.pep.*
6: /cgn2.6/prodata/2/1aa/backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	691	43.0	309	3 US-08-988-876-5	Sequence 5, Appl1
2	655	40.8	314	4 US-08-988-876-7	Sequence 7, Appl1
3	654.5	40.7	321	4 US-08-748-506-18	Sequence 18, Appl1
4	651.5	40.5	321	4 US-08-748-506-10	Sequence 10, Appl1
5	647.5	40.3	321	4 US-08-748-506-20	Sequence 20, Appl1
6	645	40.1	333	3 US-08-988-876-6	Sequence 6, Appl1
7	644.5	40.1	321	4 US-08-748-506-12	Sequence 12, Appl1
8	640	39.8	316	2 US-08-827-291A-2	Sequence 2, Appl1
9	638.5	39.7	321	4 US-08-748-506-13	Sequence 13, Appl1
10	630.5	39.2	321	4 US-08-748-506-19	Sequence 19, Appl1
11	628.5	39.1	321	4 US-08-748-506-11	Sequence 11, Appl1
12	622.5	38.7	236	2 US-08-467-948A-2	Sequence 2, Appl1
13	622.5	38.7	236	3 US-08-467-947A-2	Sequence 2, Appl1
14	596	37.1	284	1 US-08-118-270-61	Sequence 61, Appl1
15	596	37.1	284	5 PCT-US93-08528-61	Sequence 61, Appl1
16	589	37.1	327	4 US-08-748-506-24	Sequence 24, Appl1
17	589	36.7	327	4 US-08-748-506-22	Sequence 22, Appl1
18	589	36.7	327	4 US-08-748-506-13	Sequence 23, Appl1
19	579	36.0	327	4 US-08-748-506-24	Sequence 14, Appl1
20	572.5	35.6	284	1 US-08-118-270-67	Sequence 67, Appl1
21	572.5	35.6	284	5 PCT-US93-08528-67	Sequence 67, Appl1
22	548.5	34.1	293	1 US-08-118-270-60	Sequence 60, Appl1
23	548.5	34.1	293	5 PCT-US93-08528-60	Sequence 60, Appl1
24	548	34.1	247	1 US-08-465-980-3	Sequence 3, Appl1
25	548	34.1	247	2 US-09-053-303-3	Sequence 3, Appl1
26	548	34.1	247	5 PCT-US95-07093-3	Sequence 3, Appl1
27	546	34.0	277	1 US-08-118-270-62	Sequence 62, Appl1

28	546	34.0	227	5 PCT-US93-08528-62	Sequence 62, Appl1
29	529.5	32.9	286	1 US-08-118-270-65	Sequence 65, Appl1
30	529.5	32.9	286	5 PCT-US93-08528-65	Sequence 65, Appl1
31	526	32.7	273	1 US-08-118-270-63	Sequence 63, Appl1
32	526	32.7	273	5 PCT-US93-08528-63	Sequence 63, Appl1
33	525	32.7	275	1 US-08-118-270-66	Sequence 66, Appl1
34	525	32.7	275	5 PCT-US93-08528-66	Sequence 66, Appl1
35	522	32.5	277	1 US-08-118-270-68	Sequence 68, Appl1
36	522	32.5	277	5 PCT-US93-08528-68	Sequence 68, Appl1
37	520	32.4	222	2 US-08-467-948A-27	Sequence 27, Appl1
38	520	32.4	222	3 US-08-467-947A-27	Sequence 27, Appl1
39	519.5	32.3	274	1 US-08-118-270-69	Sequence 69, Appl1
40	519.5	32.3	274	5 PCT-US93-08528-69	Sequence 69, Appl1
41	519	32.3	269	1 US-08-118-270-64	Sequence 64, Appl1
42	519	32.3	269	5 PCT-US93-08528-64	Sequence 64, Appl1
43	393	24.5	320	4 US-09-439-313-527	Sequence 527, Appl1
44	377	23.5	320	1 US-08-465-980-2	Sequence 2, Appl1
45	377	23.5	320	2 US-09-053-303-2	Sequence 2, Appl1

ALIGNMENTS

RESULT 1
US-08-988-876-5
Sequence 5, Application US/08988876
Patent No. 6063596
GENERAL INFORMATION:
APPLICANT: Lal, Preeti
APPLICANT: Bandman, Olga
APPLICANT: Hillman, Jennifer L.
TITLE OF INVENTION: G PROTEIN COUPLED RECEPTORS ASSOCIATED
TITLE OF INVENTION: WITH IMMUNE RESPONSE
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/988,876
FILING DATE: Herewith
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0441 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-855-0555
TELEFAX: 650-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 309 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: 1314667
US-08-988-876-5

[illegible]

```

1      RESULT      2
2      US-08-988-876-7
3      Sequence 7, Application US/08988876
4      Patent No. 6063596
5
6      GENERAL INFORMATION:
7      APPLICANT: Lal, Preeti
8      APPLICANT: Bandman, Olga
9      APPLICANT: Hillman, Jennifer L.
10     APPLICANT: yue, Henry
11     TITLE OF INVENTION: G PROTEIN COUPLED RECEPTORS ASSOCIATED
12     TITLE OF INVENTION: WITH IMMUNE RESPONSE
13     NUMBER OF SEQUENCES: 9
14     CORRESPONDENCE ADDRESS:
15     ADDRESSEE: Inocyte Pharmaceuticals, Inc.
16     STREET: 3174 Porter Drive
17     CITY: Palo Alto
18     STATE: CA
19     COUNTRY: USA
20     ZIP: 94304
21
22     COMPUTER READABLE FORM:
23     MEDIUM TYPE: Diskette
24     COMPUTER: IBM Compatible
25     OPERATING SYSTEM: DOS
26     SOFTWARE: FASTSEQ for Windows Version 2.0
27     CURRENT APPLICATION DATA:
28     APPLICATION NUMBER: US/08/988,876
29     FILING DATE: Herewith
30     CLASSIFICATION:
31     PRIOR APPLICATION DATA:
32     APPLICATION NUMBER:
33     FILING DATE:
34     ATTORNEY/AGENT INFORMATION:
35     NAME: Billings, Lucy J.
36     REGISTRATION NUMBER: 36,749
37     REFERENCE/DOCKET NUMBER: PF-0441 US
38     TELECOMMUNICATION INFORMATION:
39     TELEPHONE: 650-855-0555
40     TELEFAX: 650-845-4166
41
42     TELEX:
43
44     INFORMATION FOR SEQ ID NO: 7:
45     SEQUENCE CHARACTERISTICS:
46     LENGTH: 314 amino acids
47     TYPE: amino acid
48     STRANDEDNESS: single
49     TOPOLOGY: linear
50
51     IMMEDIATE SOURCE:
52     LIBRARY: GenBank

```

CLONE: 32086
US-08-988-876-7

Query Match	40.8%	Score 655;	DB 3;	Length 314;
Best Local Similarity	43.5%;	Pred. No. 2.2e-50;		
Matches 131;	Conservative 61;	Mismatches 109;	Indels 0;	Gaps 0;

[illegible]

```

1      RESULT 3
2      US-08-748-506-18
3      Sequence 18, Application US/08748506
4      Patent No. 6159707
5      GENERAL INFORMATION:
6      APPLICANT: Ronnett et al.
7      TITLE OF INVENTION: NOVEL SPERM RECEPTORS
8      NUMBER OF SEQUENCES: 31
9      CORRESPONDENCE ADDRESS:
10     ADDRESSEE: Leydig, Volt & Mayer, Ltd.
11     STREET: Two Prudential Plaza, Suite 4900
12     CITY: Chicago
13     STATE: IL
14     COUNTRY: US
15     ZIP: 60601-6780
16     COMPUTER READABLE FORM:
17     MEDIUM TYPE: Floppy disk
18     COMPUTER: IBM PC compatible
19     OPERATING SYSTEM: PC-DOS/MS-DOS
20     SOFTWARE: PatentIn Release #1.0, Version #1.25
21     CURRENT APPLICATION DATA:
22     APPLICATION NUMBER: US/08/748,506
23     FILING DATE: 08-NOV-1996
24     CLASSIFICATION: 435
25     PRIOR APPLICATION DATA:
26     APPLICATION NUMBER: US 60/033,751
27     FILING DATE: 09-NOV-1995
28     CLASSIFICATION: 435
29     ATTORNEY/AGENT INFORMATION:
30     REFERENCE/DOCKET NUMBER: 74940
31     TELECOMMUNICATION INFORMATION:
32     TELEPHONE: 312-616-5600
33     TELEFAX: 312-616-5700
34     INFORMATION FOR SEQ ID NO: 18:
35     SEQUENCE CHARACTERISTICS:
36     LENGTH: 321 amino acids
37     TYPE: amino acid
38     TOPOLOGY: unknown
39     MOLECULE TYPE: protein
40     US-08-748-506-18

```

```

Query Match          40.7%; Score 654.5; DB 4; Length 321.
Best Local Similarity 44.7%; Pred. No. 2.5e-50;
Matches 135; Conservative 48; Mismatches 118; Indels 1; Gaps 1;

QY      3 WVNDSI-IQEPIILLGFSRPMLEPFLVAVPLISYVTTFGNNTILVLRSLDTRKLTMPWF 61
        | : : : | : : | : : | : : | : : | : : | : : | : : | : : |
Db       8 WOENSLIVAKHFAFKFSESVPECEFLNLMLLMLVLSLGNLLIYLALCTSPSLTPMWF 67
QY      62 FLTNISLIDLCTTCCTCYQMVLNCLSKRYSYRGCVAOQLFFLAGATEYVLLAMSFSD 121
        ||| ||||| : : : : : : : : : : : : : : : : : : : : : : : :
Db       68 FLANISLLEIGYTCSVIPKMQSIYSREAREISRRCATQMEFFEAFGECCLLAAMAFD 127
QY     122 MEVANICRPGLHYSVIMHORLIQLAAASVAWTGSNSVMNSTLTLOLPIDDPYVDIAHFLEV 181
        ||| ||| : : : : : | : | : | : : : : : : : : : : : : : :
Db     128 RCMAICSPLAHATRMSRKVCACHLAIVSMGCIVSLGOTNFIFSINFPGPCEDIHFPDDL 187
QY     182 PALKLSCVETTANEAEELFLVSELPHILPLLILISYAIFVRAVALRIOSAESGRCAFSTC 241
        ||| ||| : : : : : | : | : | : : : : | : : | : : | : : | : :
Db     188 PLLALACGDPTSQNDAEFVVAVLCISSPFLIIIVSYKKILLAVLAMPBEGRHKHALSTQC 247
QY     242 GSHLIVSLEFTSTAVSVILOPPSPBSKOQKMVSLEYGIAPMLNPITYTLNKVEKGESF 301
        ||| ||| : : : : : | : | : | : : : : | : : | : : | : : | : :
Db     248 SSHLLVLYLFYGSACTILPRKSSHSHPGMDFLAEFYVTSMLNDPIIYSRNKEVKAAAL 307
QY      302 KR 303
        :|
Db     308 RR 309

RESULT    4
US-08-748-506-10
: Sequence 10, Application US/08748506
: Patent No. 6159707
: GENERAL INFORMATION:
: APPLICANT: Ronnett et al.
: TITLE OF INVENTION: NOVEL SPERM RECEPTORS
: NUMBER OF SEQUENCES: 31
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Leydig, Volt & Mayer, Ltd.
: STREET: Two Prudential Plaza, Suite 4900
: CITY: Chicago
: STATE: IL
: COUNTRY: US
: ZIP: 60601-6780
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/748,506
: FILING DATE: 08-Nov-1996
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 60/003,751
: FILING DATE: 09-Nov-1995
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: REFERENCE/DOCKET NUMBER: 74940
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 312-616-5600
: TELEFAX: 312-616-5700
: INFORMATION FOR SEQ ID NO: 10:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 321 amino acids
: TYPE: amino acid
: STRANDEDNESS: unknown
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-08-748-506-10

```

	Query Match	40.5%;	Score 651.5;	DB 4;	Length 321.	
	Best Local Similarity	44.7%;	Pred. No. 4.6e-50;			
	Matches 135;	Conservative	47;	Mismatches 119;	Indels 1;	Gaps 1;
QY	3 WVNDSI-IQEFLILGFSRPMLEPPLLVFLSYVTITFGNLTITLVSRLDKLHTPMYE 61	:	:	:	:	:
D8	8 WOENSLFVKHFAAFKSEVPGECFLLFNLIILMFVLISLGNLTIVLAICTPSLHTPMYE 67	:	:	:	:	:
QY	62 FLTNLSLDLCYTCTTGVOMLVNLCSIRKVISYRCVAOLFFLALGATEYLILAAMSD 121	:	:	:	:	:
D8	68 FLANISLLEIGTCYSVIKRMLOSIVSEAREISRCCATQMFEFFAEFGITECCLLAMAD 127	:	:	:	:	:
QY	122 WEVAICRELHSVIMHORICCOLAAMASWVGSGNSVMWLSTLFLOLPIDPYIIDHFCEV 181	:	:	:	:	:
D8	128 RCMALCSPLMHRATRNSREVCACHALIVSMGMGCTIVSLGQTINFSLMFCGPCEIDHFDCL 187	:	:	:	:	:
QY	182 PALKLSCVETTANEADELIVSELPHLIPRTLILISYAETIVAVLRIOSAEGROKAFGTC 241	:	:	:	:	:
D8	188 PRLTALAGDPSQNEPAIFVVAVLICISSPFELLIVSVYKILLIIVALLMPSPREGRHKLSTC 247	:	:	:	:	:
QY	242 GSHLIVLSLFYSTANSVVLQPPSPSSKDGKNVSLFYGIILAMPILNPITYLANKVEKEEF 301	:	:	:	:	:
D8	248 SSHLLVTLVFYGSACTILRPKSSHSPGMDKFLAELFYVTSTMNPDIYSLNKKEYKAAAL 307	:	:	:	:	:
QY	302 KR 303	:	:	:	:	:
D8	308 RR 309	:	:	:	:	:

RESULT 5
 US-08-748-506-20
 : Sequence 20, Application US/08748506
 : Patent No. 6159707
 : GENERAL INFORMATION:
 : APPLICANT: Romnett et al.
 : TITLE OF INVENTION: NOVEL SPERM RECEPTORS
 : NUMBER OF SEQUENCES: 31
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: Leydig, Voigt & Mayer, Ltd.
 : STREET: Two Prudential Plaza, Suite 4900
 : CITY: Chicago
 : STATE: IL
 : COUNTRY: US
 : ZIP: 60601-6780
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Floppy disk
 : COMPUTER: IBM PC compatible
 : OPERATING SYSTEM: PC-DOS/MS-DOS
 : SOFTWARE: Patentin Release #1.0, Version #1.25
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/08/748,506
 : FILING DATE: 08-NOV-1996
 : CLASSIFICATION: 435
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: US 60/033,751
 : FILING DATE: 09-NOV-1995
 : CLASSIFICATION: 435
 : ATTORNEY/AGENT INFORMATION: 74940
 : REFERENCE/DOCKET NUMBER:
 : TELECOMMUNICATION INFORMATION:
 : TELEPHONE: 312-616-5600
 : TELEFAX: 312-616-5700
 : INFORMATION FOR SEQ ID NO: 20:
 : SEQUENCE CHARACTERISTICS:
 : LENGTH: 321 amino acids
 : TYPE: amino acid
 : TOPOLOGY: unknown
 : MOLECULE TYPE: protein
 : US-08-748-506-20

Best Local Similarity 42.7%; Pred. No. 1e-49;
Matches 134; Conservative 56; Mismatches 121; Indels 3; Gaps 2;

QY	3	WVNDSI--IOEFLILGEGSRPMLEPPLVAVLIVSTVYIFGVLITIVSRDTRKHTPMYF	61
Db	8	WOENLSYKRRFAKFSFVSVECFLLFTLLILMTLVSTGMIILALACISPSLHTMYF	67
QY	62	FLTNLSLIDLCYTCYVPMQLVNLCSIRKVISYGCYAQLFIFALGATEYLLAVMSFD	121
Db	68	FLANLSLEIEYTCYSVPKMLQSLVSEARGSLMSGCASOMEFFEFGTECCLLAAMAFD	127
QY	122	WEFAICRPLHYSVIMHQLCOLLAASAVWTFSSWMLSTLTLDLPICDPAVIDHFCSEV	181
Db	128	RYMAICSPHLHATNRSKRCVAKYLAIVSMVCCIVGLQOTINIFSLNCGCPREIDHFFCDL	187
QY	182	PALKLKSCVETTANDELELVSELPHLPLTLTLIISAFYRAVLRIQSAEGROKAFGTC	241
Db	188	PPLALACGDSQNEAALFYAAVLCIFSPFLLIISVYRIIVAVLWMPSPGGRKALSTGC	247
QY	242	GSHLIIIVLFSTAVSYVLQPPSPSSKQOGKMSLEFYGIAPMNPILYTLRNKEVEGF	301
Db	248	SSHLLVYVLEFGYSNATYLRKSSKSHSPGVCKLLDLFTTSVSMNPILIISLRNKEVGAL	307
QY	302	KRLVA--RVFLIKK 313	
Db	308	RTTGELKKVLTMRK 321	

RESULT 6
US-08-988-876-6

GENERAL INFORMATION:
APPLICANT: Lal, Preeti
APPLICANT: Bandman, Olga
APPLICANT: Hillman, Jennifer L.
APPLICANT: Yue, Henry
TITLE OF INVENTION: G PROTEIN COUPLED RECEPTORS ASSOCIATED
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/988,876
FILING DATE: Herewith
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0441 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-855-0555
TELEFAX: 650-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 333 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear

;
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 205814
US-08-988-876-6

Query Match	40.1%	Score 645	DB 3	Length 333
Best Local Similarity	40.9%	Pred. No. 1.8e-49		
Matches 124	Conservative 62	Mismatches 117	Indels 0	Gaps 0

[illegible]

RESULT 7
US-08-748-506-12

```

1  GENERAL INFORMATION:
2  APPLICANT: Ronnelt et al.
3  TITLE OF INVENTION: NOVEL SPERM RECEPTORS
4  NUMBER OF SEQUENCES: 31
5  CORRESPONDENCE ADDRESS:
6  ADDRESSEE: Leydig, Volt & Mayer, Ltd.
7  STREET: Two Prudential Plaza, Suite 4900
8  CITY: Chicago
9  STATE: IL
10 COUNTRY: US
11 ZIP: 60601-6780
12 COMPUTER READABLE FORM:
13 MEDIUM TYPE: Floppy disk
14 COMPUTER: IBM PC compatible
15 OPERATING SYSTEM: PC-DOS/MS-DOS
16 SOFTWARE: Patentin Release #1.0, Version #1.25
17 CURRENT APPLICATION DATA:
18 APPLICATION NUMBER: US/08/748,506
19 FILING DATE: 08-NOV-1996
20 CLASSIFICATION: 435
21 PRIOR APPLICATION DATA:
22 APPLICATION NUMBER: US 60/033,751
23 FILING DATE: 09-NOV-1995
24 CLASSIFICATION: 435
25 ATTORNEY/AGENT INFORMATION:
26 REFERENCE/DOCKET NUMBER: 74940
27 TELECOMMUNICATION INFORMATION:
28 TELEPHONE: 312-616-5700
29 TELEFAX: 312-616-5700
30 INFORMATION FOR SEQ ID NO: 12:
31 SEQUENCE CHARACTERISTICS:
32 LENGTH: 321 amino acids
33 TYPE: amino acid
34 TOPOLOGY: linear

```


MOLECULE TYPE: protein
US-08-748-506-12

Query Match	40.1%;	Score 644.5;	DB 4;	Length 321;
Best Local Similarity	-42.7%;	Pred. No. 1.9e-49;		
Matches 134; Conservative	55;	Mismatches 122;	Indels 3;	Gaps 2;

Qy	3	WVNDIS-IOEIIILGFSDRPMEPLLVLLISTVYIPDNNTIILVSRDITKHPMYE	61
	8	WQENSLSVKRAFAKESFVEGECEFLFTLLILMEVLSYNTNTLLALICTSPSLHPMYE	67
Qy	62	FLTNMLSLDLCYTCYCPOMLVNVCISIRKYSYRGCAOLEIFPLGATGEYLLAVMSFD	120
Dd	68	FLNALSLIEIGYTSVYIPKMIQSVSARGISWEGCASQMEFFIFEGSITCCLLANAFD	127
Qy	122	WFAVACRPLHYSVYIMHORLCIQLAAASWYGFNSVWLSTLTQLPLCDEYVIDHELCFV	181
Dd	128	RYMAGCPHATKMSNGVCAYLIAYSWNGCIYGLQGTNFIISLNCNGCFEIDHFECDL	187
Qy	182	PALLIKISCVETTANAEALFVSELFHILIPLLIILISYAFIVRAVLRIQSAEGROKAGTC	241
Dd	188	PRLTALACGPTSQNEAALFVAAYLCIPSEFLIITSSYRILVAVLVMPSEGRHKLSTC	247
Qy	242	GSHLIIVSLFYSTAVSYVYLDPPSSSKDQKMSLFGITAIAPMLNPLIYLRNKEVEGF	300
Dd	248	SSHLLVYVTFEYGSTSAVYLRKSSSHSGVCKLLALFYTSTMLNPLIYSLRNKEVGAL	307
Qy	302	KRLVA--RVFLIKK	313
Dd	308	KRTIGLKKVLTMKR	321

US-08-827-291A-2
Sequence 2, Application US/08827291A
Patent No. 5874243
GENERAL INFORMATION:
APPLICANT: Macina, Roberto
APPLICANT: Sathe, Ganesh
TITLE OF INVENTION: NOVEL OLRC15 RECEPTOR
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: SmithKline Beecham Corporation
STREET: 709 Swedeland Road
CITY: King of Prussia
STATE: PA
COUNTRY:
ZIP: 19406
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/827,291A
FILING DATE: 28-MAR-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: King, William T
REGISTRATION NUMBER: 30,954
REFERENCE/DOCKET NUMBER: GP50001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-5015
TELEFAX: 610-270-5090
TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 316 amino acids
TYPE: amino acid

```

; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-827-291A-2

```

Query Match	39.88;	Score 640;	DB 2;	Length 316;
Best Local Similarity	42.9%;	Pred. No. 4.7e-49;		
Matches 129;	Conservative 57;	Mismatches 113;	Indels 2;	Gaps 2;

[illegible]

US-08-748-506-13
 9
 Sequence 13, Application US/08748506
 Patent No. 6135707
 GENERAL INFORMATION:
 APPLICANT: Ronnett et al.
 TITLE OF INVENTION: NOVEL SPERM RECEPTORS
 NUMBER OF SEQUENCES: 31
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Leydig, Volt & Mayer, Ltd.
 STREET: Two Prudential Plaza, Suite 4900
 City: Chicago
 STATE: IL
 COUNTRY: US
 ZIP: 60601-6780
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/748,506
 FILING DATE: 08-NOV-1996
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 60/033,751
 FILING DATE: 09-NOV-1995
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 REFERENCE/DOCKET NUMBER: 74940
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 312-616-5600
 TELEFAX: 312-616-5700
 INFORMATION FOR SEO ID NO: 13:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 321 amino acids
 type: amino acid
 STRANDEDNESS: unknown

; MOLECULE TYPE: protein

US-08-748-506-11

TELEFAX: 202-371-2540
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 296 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-467-947A-2

Query Match 38.7%; Score 622.5; DB 3; Length 296;
Best Local Similarity 43.3%; Pred. No. 1.5e-47;
Matches 123; Conservative 62; Mismatches 98; Indels 1; Gaps 1;

QY 5 NDSIIQFILLGFSRDMLEFPLLVFLISYVYIFGNLIILVSRDLKHPMEFELT 64
D 8 NQWVTEFLILGFLGRIGRIQLFGLSLFVYVFLNGITLGLISDSRLHTPMYFELS 67
QY 65 NLSLIDLCYTCYTPQMLVNLCSIRKVIYIRGCYAOLEFPLALGATEYLLLVASFTMEV 124
D 68 HLAIVNATVACNTVPMQMLVNLHPAKPIISFAGCMTDPLFLSFATTECLLVMSYDRYV 127
QY 125 AICRPLHYIYIMHQRCLQIAASWVGFNSVWLTSTLQPLCDPYVIDHFLCEVPAL 184
D 128 AICRPLHYIYIMHQRCLQIAASWVGFNSVWLTSTLQPLCDPYVIDHFLCEVPAL 187
QY 185 LKSCVETANAELEFLVSEFLHPLIPLILISYAFIVRAVLRIQSAEGROKAFGTGSH 244
D 188 LRLACADTWLNQVYIFPACMEILVGLVLSYSHTLGLTIRIQSGEGRKARSTSSH 247
QY 245 LIYVSLEFYSTAVSYILOPPSPSSKDGKMSLEFGITA-PMANP 267
D 248 LCVAGLFGSALVYMAKSRHPEQOKVLFLIQLSTPLMKP 291

RESULT 14
US-08-118-270-61
; Sequence 61, Application US/08118270
; Patent No. 5508384
; GENERAL INFORMATION:
; APPLICANT: Murphy, Randall B.
; APPLICANT: Schuster, David I.
; TITLE OF INVENTION: POLYPEPTIDES OF G-COUPLED PROTEIN
; TITLE OF INVENTION: RECEPTORS, AND COMPOSITIONS AND METHODS THEREOF
; NUMBER OF SEQUENCES: 348
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK
; STREET: 419 Seventh Street, N.W., Suite 300
; City: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: IBM floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/118,270
; FILING DATE: 09-SEP-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/943,236
; FILING DATE: 10-SEP-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Townsend, Kevin G.
; REGISTRATION NUMBER: 34,033
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528
; TELEX: 248633
; INFORMATION FOR SEQ ID NO: 61:
; SEQUENCE CHARACTERISTICS:

LENGTH: 284 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-118-270-61

Query Match 37.1%; Score 596; DB 1; Length 284;
Best Local Similarity 41.5%; Pred. No. 3.1e-45;
Matches 119; Conservative 52; Mismatches 112; Indels 4; Gaps 2;

QY 27 LVAFLISYVYITIGNFTIIVSRDLKHTPMYFFFNLSLDCYTCYTPQMLVNL 86
D 2 LFLFLIMYATVIGNFTIIVSRDLKHTPMYFFFNLSLDCYTCYTPQMLVNL 89
QY 87 SIRKVIYIRGCYAOLEFPLALGATEYLLLVASFTMEVVAICRPLHYIYIMHQRCLQIAA 146
D 62 LGSQALISFGSLTQYFLVAFGNMDFLAVMSYDRYVAILCHPLHYTTI--KQLCVLLV 119
QY 147 ASWVGFNSVWLTSTLQPLCDPYVIDHFLCEVPALILKSCVETANAELEFLVSELE 206
D 120 GSWVAMNCLHILLIMARKSFAD--LPHFCDGTPLKLKSCDTHLNLMLTEGAVY 177
QY 207 HLEPLILISYAFIVRAVLRIQSAEGROKAFGTGSHLIYVSLEFYSTAVSYILOPPSPS 266
D 178 MYPFVCLILSYIHITCAVLKVSPPRGKWSFTGSHIAYVCLFYGTVLAVYENPSSH 237
QY 267 SKDGKMSLEFGITA-PMANPILVITLANKKEYGEFKRLVAVFLIKK 313
D 238 LAGDMAAAVAVYVTPMINPFTSLKNSDKALKRLVLAIRFESKQ 284

RESULT 15
PCT-US93-08528-61
; Sequence 61, Application PC/TUS9308528
; GENERAL INFORMATION:
; APPLICANT: New York University
; TITLE OF INVENTION: POLYPEPTIDES OF G-COUPLED PROTEIN
; TITLE OF INVENTION: RECEPTORS, AND COMPOSITIONS AND METHODS THEREOF
; NUMBER OF SEQUENCES: 348
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK
; STREET: 419 Seventh Street, N.W., Suite 300
; City: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/08528
; FILING DATE: 09-SEP-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/943,236
; FILING DATE: 10-SEP-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Townsend, Kevin G.
; REGISTRATION NUMBER: 34,033
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528
; TELEX: 248633
; INFORMATION FOR SEQ ID NO: 61:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 284 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd

OM protein - protein search, using sw model

Run on: August 25, 2002, 19:02:17 ; Search time 47.87 Seconds
(without alignments)
628.284 Million cell updates/sec

Title:	US-09-800-321A-4
Perfect score:	1607
Sequence:	1 MMVNVDSIIQEIFILGFSDR.....NKEYEGFKRLVARVELIKK 313

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

```
Searched:      283138 seqs, 96089334 residues
Total number of hits satisfying chosen parameters: 283138
```

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing:	Minimum Match 0%
	Maximum Match 100%
	Listing first 45 summaries

```
Database : PLR_11:
1: plr1:*
2: plr2:*
3: plr3:*
4: plr4:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	902	56.1	312	2	A46247	olfactory receptor
2	851	53.0	316	2	A57069	olfactory receptor
3	747	46.5	315	2	JC5836	olfactory receptor
4	728	45.3	305	2	S29711	olfactory factor C
5	701	43.6	309	1	S51356	olfactory receptor
6	697	43.4	304	2	S29709	olfactory receptor
7	680	42.3	314	2	S29707	olfactory receptor
8	674.5	42.0	313	2	S20571	olfactory receptor
9	669	41.6	313	2	B23701	olfactory receptor
10	668.5	41.6	318	2	JC5202	chemoreceptor TB64
11	661	41.1	315	2	JC4658	olfactory receptor
12	660	41.1	319	2	JC5624	olfactory receptor
13	659	41.0	327	2	F23701	olfactory receptor
14	656	40.8	314	2	H23701	olfactory receptor
15	655	40.8	314	2	S20572	olfactory receptor
16	648.5	40.4	320	2	S20573	olfactory receptor
17	648	40.3	310	2	E23701	olfactory receptor
18	647	40.3	314	2	A37286	olfactory receptor
19	645	40.1	333	2	A23701	olfactory receptor
20	640	39.8	315	2	JC5201	chemoreceptor TB5650
21	640	39.8	317	2	D23701	olfactory receptor
22	630	39.2	312	2	G23701	olfactory receptor
23	628	39.1	311	2	JC5200	chemoreceptor TB3308
24	627.5	39.0	312	2	S29708	olfactory receptor
25	626.5	39.0	312	2	A48413	probable olfactory
26	623	38.8	312	2	I23701	olfactory receptor
27	611.5	38.1	307	2	S29710	olfactory receptor
28	602	37.5	311	2	C23701	olfactory receptor
29	570	35.5	312	2	A46750	olfactory receptor

30	527	32.8	264	2	PC4369	olfactory receptor
31	520	32.4	222	2	DA0745	odorant receptor (
32	513	31.9	232	2	S29001	G protein-coupled
33	477	29.7	234	2	S29000	protein-coupled
34	462	28.7	157	2	S58035	probable olfactory
35	455	28.3	216	2	I38480	olfactory receptor
36	450	28.0	222	2	B40745	odorant receptor (
37	448	27.9	216	2	I38470	olfactory receptor
38	444	27.6	216	2	I38474	olfactory receptor
39	444	27.6	234	2	S28998	G protein-coupled
40	441	27.4	234	2	S28999	G protein-coupled
41	440	27.4	216	2	I38481	olfactory receptor
42	437.5	27.2	225	2	I38478	olfactory receptor
43	434	27.0	216	2	I38479	olfactory receptor
44	434	27.0	216	2	I38476	olfactory receptor
45	423	26.3	216	2	I38477	olfactory receptor

ALIGNMENTS

```

RESULT      1
A46247
olfactory receptor OR3 - mouse
C:Species: Mus musculus (house mouse)
C:Date: 22-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 26-Aug-1999
C:Accession: A46247
R:Nef, P.; Hermans-Borgmeyer, I.; Arterios-Pin, H.; Beasley, L.; Dlonne, V.E.; Heinemann, Proc. Natl. Acad. Sci. U.S.A. 89, 8948-8952, 1992
A:Title: Spatial pattern of receptor expression in the olfactory epithelium.
A:Reference number: A46247; MUID:93028384
A:Accession: A46247
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: nucleic acid
A:Residues: 1-312 <NEP>
A:Cross-references: GB:864005; NID:g200153; PIDN:AAA39682.1; PID:g200154
A>Note: sequence extracted from NCBI backbone (NCBIP:115362)
C:Superfamily: olfactory receptor OR14

Query Match          56.1%; Score 902; DB 2; Length 312;
Best Local Similarity 58.8%; Pred. No. 1.3e-72;
Matches 174; Conservative 46; Mismatches 76; Indels 0; Gaps 0;

Oy   12  FLTGFSRPMLEFPLVFLISTVTTFGNLTITLVSRLDTKLTHTPWFFELTNLSLDL  71
    |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db   12  FILMGVDSPHLEITEFAVLIASTYLLTVGNLTITLISRLDARLHTPWFFELSNSSIDL  71

Oy   72  CYTCTGVQMLVNLCSIRKVTSYRCVAQLFTFLAGATEYLLLAVMSEDFWVAICRPLH  131
    ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db   72  AFTTSSVQMOKNLIMGPKRTIYSTGCYQLVFLMGLGTECLLVMAFDKYAVACRPDLH  131

Oy   132 YSVIMHQRIQLAAASWVGTSNSWMLSTTLQLPLCDPVYDIHFCEVPALTKSCVE  191
    |::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db   132 YMTVNPRLCMGLAIASWLGIGNSVIQTFTFLQLPFCGHKKRVDNFLCEVPAMIKLAGSD  191

Oy   192 TTYANEAEFLVSELFLPLILISYAFIVRAVLRIQSAEGROKAFGTCGSHLYVSLF  251
    |::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db   192 TSLNEAVINGCVETFTVPVPSVILVSYCFIQADVKIRSVESGRRAFTMCVSHLEVLF  251

Oy   252 YSTAASVYLQPPSPSKDOGMVSLFYGIAMPNLNPLTYLTRNKREVRGEFRIVAR  307
    |::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db   252 YGSATYGLLPKASSNQSGKFISLIFYSVYPMVNPPLTYLTRNKREVRKALGRLLQG  307

RESULT      2
A57069
olfactory receptor FAT11 - human
C:Species: Homo sapiens (man)
C:Date: 03-Oct-1995 #sequence_revision 03-Oct-1995 #text_change 26-Aug-1999
C:Accession: A57069
R:Fan, W.; Liu, Y.C.; Parimoo, S.; Weissman, S.M.
Genomics 27, 119-123, 1995

```

A:Title: olfactory receptor-like genes are located in the human major histocompatibility complex
A:Reference number: A57069, MUID:95394447
A:Accession: A57069
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-316 <FAN>
A:Cross-references: GB:I35475; NID:g1041044; PIDN:AAB36567.1; PID:g601919
C:Genetics:
C:Gene: GDB:FAN11; OLFR2
A:Cross-references: GDB:I323249; OMIM:600578
A:Map position: 6p21.3-6p21.3
C:Superfamily: olfactory receptor OR14

[illegible]

RESULT 3
JC5836
Olfactory receptor-like protein 2 - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 05-Mar-1998 #sequence_revision 13-Mar-1998 #text_change 24-Nov-1999
C:Accession: J05836
R:Blache, P.; Gros, L.; Salazar, G.; Bataille, D.
Biochem. Biophys. Res. Commun. 242, 669-672, 1998
A:Title: Cloning and tissue distribution of a new rat olfactory receptor-like
A:Reference number: J05836; M01D:9812531
A:Accession: JC5836
A:Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-315 <BLA>
A:Cross-references: GB:AF029357; NID:g2570934; PIDN:AAC3969.1; PID:g2570935
C:Comment: This protein is implicated in white cell maturation and migration.
C:Superfamily: olfactory receptor OR14
F:64-86/Domain: transmembrane #status predicted <TM1>
F:128-53/Domain: transmembrane #status predicted <TM2>
F:95-123/Domain: transmembrane #status predicted <TM3>
F:147-165/Domain: transmembrane #status predicted <TM4>
F:203-229/Domain: transmembrane #status predicted <TM5>
F:241-264/Domain: transmembrane #status predicted <TM6>
F:277-293/Domain: transmembrane #status predicted <TM7>

Query Match	46.5%;	Score 747;	DB 2;	Length 315;
Best Local Similarity	46.2%;	Pred. No. 7.8e-59;		
Matches 139;	Conservative 71;	Mismatches 91;	Indels 0;	Gaps 0;

QY 5 NDSQQEELLGSGRPLEFPLVFLIVSTVFPGVLITILSRDLKHTPEMFEFT 64
| | | | | | | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 8 NCSSVSSEFIILGFESSESQIRMALEFFELLLWTLVGLGLVALITYDLSRHHMTFFELS 67

QY 65 NLSLDDLCYTTCVPQMVLNCLSIKRVISYRGCAOALEIFPALGATEYLALVAVSFDMEV 124
| | | | : | | | | | | | | : | : | : | | | | : | : | : | : | : | : | : | : | : | :
Db 68 ILSLVDMSGYTTTYPQMLVNMYCPRRTISMGACVAAQMEFILVLDAIECVLAIAWAYRYI 127

QY 125 AICRPLRLSVIMHQRLCQLAASAWVTGFSNSVSLSTLTLOPLCDPYVIDHFLCEVPAL 184
| | | | | | | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 128 AICEFLPHXSVLMRSRLVCAMWTICSSISVTGALLTYFTFMRLPCGCPKIINHFCEVPAV 187

QY 165 LKLSCVEETANDELFVLSELPHLIPLLLIIISAFTAVLRATLRQSAGROKARCTGSGH 244
| | | | : | : | : | : | : | : | | | | | | | | : | : | : | : | : | : | : | : | : | :
Db 188 LKLACADITSFNDRIOLFVLLLPISLIASAACIFVSLTRRSSOGRLKSTCASH 247

QY 245 LIVVELFSTAVSYVLQPPSPSKDGCKMVASLFGIAPMLNPILTYLTRNKVEVEGCKRL 304
| | | : | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 248 ITVTVMFTGPAMVMTMRGSMWDPERDKKLLFNIVVSAFLNPILTIYSLNKNDVGAKMKV 307

QY 305 V 305
:
Db 308 L 308

RESULT 4
S29711
olfactory factor OR37 - rat
C.Species: Rattus norvegicus (Norway rat)
C.Date: 25-Feb-1994 #sequence_revision 10-Nov-1995 #text_change 04-Sep-1998
C.Accession: S29711
R.Ramung, K.; Krüger, J.; Strotmann, J.; Boehhoff, I.; Kubick, S.; Baumstark, C.; Br
A.Title: Cloning and expression of odorant receptors.
A.Reference number: S29707; MUID:93149273
A.Accession: S29711
A.Molecule type: mRNA
A.Residues: 1-305 <RAW>
C:Superfamily: olfactory receptor OR14

[illegible]

RESULT 5
S51356
olfactory receptor - rat

Db 245 LSVSLFSGTVIPLYLCPSSNNSTYKETFVMSLMTYLVTPMLNPFIFYSLRRNDIK 298

RESULT 8

S20571

olfactory receptor - dog

C:Species: Canis lupus familiaris (dog)

C:Date: 22-Nov-1993 #sequence_revision 01-Dec-1995 #text_change 26-Aug-1999

C:Accession: S20571

R:Pamrentliar, M.; Libert, F.; Schumans, S.; Schifmann, S.; Lefort, A.; Eggerickx, D.; Nature 355, 453-455, 1992

A:Title: Expression of members of the putative olfactory receptor gene family in mammals

A:Reference number: S20571; MUID:92131132

A:Accession: S20571

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1513 <PAR>

A:Cross-references: EMBL:X64966; NID:8890; PTDN:CAA46129.1; PID:8891

C:Superfamily: olfactory receptor OR14

C:Keywords: G protein-coupled receptor; transmembrane protein

Query Match	42.0%;	Score 674.5;	DB 2;	Length 313;
Best Local Similarly	43.9%;	Pred. No. 2.1e-52;		
Matches 133; Conservative	59;	Mismatches 110;	Indels 1;	Gaps 1;

[illegible]

```

RESULT      9
B23701
Olfactory receptor F5 - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 22-Jan-1993 #sequence_revision 04-Sep-1998
C:Accession: B23701
R:Buck, L.; Axel, R.
Cell 65, 175-187, 1991
A:Title: A novel multigene family may encode odorant receptors: a molecular basis for co
A:Reference number: A23701; MUID:91191556
A:Accession: B23701
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-313 <HUC>
A:Cross-references: GB:M64377
C:Superfamily: olfactory receptor OR14
C:Keywords: G protein-coupled receptor; transmembrane protein

```

Query Match	41.6%	Score 669	DB 2	Length 313
Best Local Similarity	42.8%	Pred. No. 6	5e-52	
Matches 134	Conservative	53	Mismatches 126	Indels 0
			Gaps	0

[illegible]

RESULT 10

JC5202

Chemoreceptor TB641 - rat

C:Species: Rattus norvegicus (Norway rat)

C:Date: 20-Feb-1997 #sequence_revision 27-Feb-1997 #text-change 26-Aug-1999

C:Accession: JC5202; PC4304

R:Thomas, M.B.; Haines, S.L.; Akesson, R.A.

Gene 178, 1-5, 1996

A:Title: Chemoreceptors expressed in taste, olfactory and male reproductive tissues.

A:Reference number: JC5200; MUID:97080538

A:Accession: JC5202

A:Status: preliminary; nucleic acid sequence not shown

A:Molecule type: DNA

A:Residues: 1-318 <TH01>

A:Cross-references: GH:U0949; NID:g1256392; PIDN:AAC52911.1; PID:g1256393

A:Accession: PC4304

A:Status: preliminary

A:Molecule type: protein

A:Residues: 145-153; 245-253 <TH02>

A:Experimental source: taste bud

C:Comment: This protein is coupled to a GTP-binding protein-mediated signal transduct

C:Genetics:

A:Gene: tb641

A:Superfamily: olfactory receptor OR14

C:Keywords: olfaction; taste bud; transmembrane protein

F:30-53/Domain: transmembrane #status predicted <TM1>

F:63-84/Domain: transmembrane #status predicted <TM2>

F:106-125/Domain: transmembrane #status predicted <TM3>

F:145-159/Domain: transmembrane #status predicted <TM4>

F:202-224/Domain: transmembrane #status predicted <TM5>

F:243-265/Domain: transmembrane #status predicted <TM6>

F:276-296/Domain: transmembrane #status predicted <TM7>

Query Match	41.6%;	Score 668.5;	DB 2;	Length 318;
Best Local Similarity	44.08;	Pred. No. 7.3e-52;		
Matches 133; Conservative	65;	Mismatches 101;	Indels 3;	Gaps 2;

QY 6 DS.OE.IELFGE.SRPM.EEPFLVYV.FLSTAVTE.EGML.IIIVS.RD.KT.KHT - PM.FE.LT 64
10 D.VYVTD.FELL.LG.LA.RPN.RK.FLE.LV.E.LL.IY.IY.IY.Q.G.M.L.L.L.L.L.L.L.Y.V.A.D.P.K.H.A.R.M.Y.L.L.L.G 69
QY 65 N.L.S.L.D.I.C.Y.T.T.C.Y.V.P.O.M.L.V.N.C.S.T.R.K.Y.I.S.T.R.G.C.V.A.O.L.F.E.L.A.G.A.T.E.Y.U.L.L.A.V.S.D.E.W.E.V 124
Db 70 V.L.S.E.L.D.M.W.L.S.S.V.I.P.R.I.I.N.E.P.P.A.N.K.A.I.A.R.G.C.G.A.O.L.Y.E.F.F.H.G.S.T.O.C.F.L.Y.T.T.A.A.V.D.R.Y.L 129
QY 125 A.I.C.P.E.H.Y.S.V.T.M.H.O.R.C.I.O.L.A.A.S.W.V.G.F.S.N.V.L.S.T.L.T.O.L.P.L.C.D.E.V.V.I.D.H.E.C.E.V.P.A.L 184
Y

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 25, 2002, 20:08:02 ; Search time 37.38 Seconds

(without alignments)
324.217 Million cell updates/sec

Title: US-09-800-321A-4

Perfect score: 1607

Sequence: 1 MNVYNDSTIIEFILLGFSDR.....NKEVKEGKRLVANFLIKK 313

Scoring table: BIOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 08

Maximum Match 1008

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1593	99.1	313	1	O2B6_HUMAN
2	1295	80.6	357	1	O2B7_HUMAN
3	1130	70.3	313	1	O2B3_HUMAN
4	933	58.1	311	1	O2J3_HUMAN
5	904	56.3	312	1	O2U2_HUMAN
6	902	56.1	312	1	OLJ5_MOUSE
7	889	55.3	312	1	O2H2_MOUSE
8	885	55.1	316	1	O2H3_HUMAN
9	883	54.9	312	1	O2C1_HUMAN
10	871	54.2	320	1	O2M1_HUMAN
11	833	53.1	316	1	O2H1_HUMAN
12	804	50.0	317	1	OLF3_CANFA
13	779	48.5	317	1	O2F1_HUMAN
14	767	47.7	321	1	O5V1_HUMAN
15	766	47.7	317	1	O2E2_HUMAN
16	757	47.1	308	1	O2D2_HUMAN
17	711.5	44.3	309	1	O7A1_HUMAN
18	701	43.6	313	1	O1C1_HUMAN
19	697	43.4	319	1	O2S2_HUMAN
20	693.5	43.2	315	1	OAA4_HUMAN
21	691	43.0	309	1	OLR4_CANFA
22	688	42.8	314	1	O5F1_HUMAN
23	687	42.8	314	1	O5I1_HUMAN
24	687	42.8	320	1	O7C1_HUMAN
25	686	42.7	309	1	O7AH_HUMAN
26	683	42.5	312	1	OAC1_HUMAN
27	681	42.4	311	1	OLR2_CANFA
28	681	42.4	320	1	O7C2_HUMAN
29	680	42.3	311	1	OLF1_CANFA
30	680	42.3	312	1	OLF4_CHICK
31	679	42.3	313	1	O1G1_HUMAN
32	676	42.1	312	1	O1F1_HUMAN
33	675	42.0	313	1	OLF5_RAT

34	674.5	42.0	313	1	OLR2_CANFA	P30955 canis faml
35	669	41.6	311	1	O8B8_HUMAN	Q15620 homo sapien
36	668.5	41.6	310	1	O4D1_HUMAN	Q15615 homo sapien
37	668	41.6	317	1	OAA5_HUMAN	Q9H207 homo sapien
38	665	41.4	314	1	OAA3_HUMAN	P58181 homo sapien
39	663	41.3	318	1	OLR1_CHICK	P37067 gallus gall
40	662	41.2	310	1	O2A4_HUMAN	Q95047 homo sapien
41	661	41.1	315	1	O3A1_HUMAN	P47881 homo sapien
42	660.5	41.1	307	1	OCB2_HUMAN	P58182 homo sapien
43	660	41.1	319	1	O7A5_HUMAN	Q15622 homo sapien
44	659	41.0	327	1	OLF7_RAT	P23270 rattus norv
45	658	40.9	312	1	OLF5_CHICK	P37071 gallus gall

ALIGNMENTS

RESULT	ID	STANDARD	PRT	AA
1	O2B6_HUMAN			
1	O2B6_HUMAN			
AC	P58173; Q9H5B0;			
DT	16-OCT-2001 (Rel. 40, Created)			
DT	16-OCT-2001 (Rel. 40, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Olfactory receptor 286 (hs6M1-32) (Olfactory receptor 6-31) (OR6-31).			
GN	OR2B6.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Williams S.;			
RL	Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.			
CC	-1- FUNCTION: PUTATIVE ODORANT RECEPTOR.			
CC	-1- SUBCELLULAR LOCATION: Integral membrane protein.			
CC	-1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; AL133267; CAC14158.1;			
DR	InterPro; IPR000276; GPCR_Rhodopsn.			
DR	Pfam; PF00001; 7tm.1.1.			
DR	PRINTS; PR00237; GPCR_RHODOPS.			
DR	PROSITE; PS00237; G-PROTEIN_RECP_FL1.1.			
DR	PROSITE; PS0262; G-PROTEIN_RECP_FL2.1.			
KW	G-protein coupled receptor; Transmembrane; Glycoprotein;			
KW	Multigene family; Olfaction.			
FT	DOMAIN 1 25			
FT	TRANSMEM 26 49			
FT	DOMAIN 50 57			
FT	TRANSMEM 58 79			
FT	DOMAIN 80 100			
FT	TRANSMEM 101 120			
FT	DOMAIN 121 139			
FT	TRANSMEM 140 158			
FT	DOMAIN 159 195			
FT	TRANSMEM 196 219			
FT	DOMAIN 220 236			
FT	TRANSMEM 237 259			
FT	DOMAIN 260 272			
FT	TRANSMEM 273 292			
FT	DOMAIN 293 313			
FT	DISULFID 97 189			
FT	CARBOHYD 5 5			
SQ	SEQUENCE 313 AA; 35414 MW; 71DA59541ACF5501 CRC64;			

Query Match 99.1%; Score 1593; DB 1; Length 313;
Best Local Similarity 99.7%; Pred. No. 3, 9e-112;
Matches 312; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MNVNVDSIIIOEFLILGSDRPMLEFPLVYFLISYVTITGNTITLIVSRIDTKLHPMY 60
DB 1 MNVNVDSIIIOEFLILGSDRPMLEFPLVYFLISYVTITGNTITLIVSRIDTKLHPMY 60

QY 61 FFLTNLSLDLCYTTCTVPMOLVNLCSIRKVISYRGCAVOLFPLALGATEYLLAVMSF 120
DB 61 FFLTNLSLDLCYTTCTVPMOLVNLCSIRKVISYRGCAVOLFPLALGATEYLLAVMSF 120

QY 121 DMFVAICRPLHYSVIMHQRICLOLAASVWTGFSNSVWSTLTLOLPDCPYVIDHFCE 180
DB 121 DMFVAICRPLHYSVIMHQRICLOLAASVWTGFSNSVWSTLTLOLPDCPYVIDHFCE 180

QY 181 VPALKLISCVEETANAEELFLVSEFLPLILPLILISYATVAVARIQSAEGRQKAFGT 240
DB 181 VPALKLISCVEETANAEELFLVSEFLPLILPLILISYATVAVARIQSAEGRQKAFGT 240

QY 241 CGSHLIVSLFSTAVSVYIQPPSPSSKDGKAVSLFYGIAPMLNPLIYTLRNKEVKEG 300
DB 241 CGSHLIVSLFSTAVSVYIQPPSPSSKDGKAVSLFYGIAPMLNPLIYTLRNKEVKEG 300

QY 301 FKRIVARVFLIKK 313
DB 301 FKRIVARVFLIKK 313

RESULT 2
02B2_HUMAN STANDARD; PRT; 357 AA.
AC Q9GZK3; Q9GZL2; Q9Y299;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Olfactory receptor 2B2 (Olfactory receptor 6-1) (OR6-1) (Hs6M1-10).
GN OR2B2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND VARIANT ARG-234.
RA Ziegler A., Ehlers A., Forbes S.A., Trowsdale J., Uchanska-Ziegler B.,
RA Volz A., Younger R., Beck S.;
RT "Polymorphic olfactory receptor genes and HLA loci constitute extended
haplotypes";
RT (in) Kasahara M. (eds.);
RL Major histocompatibility complex-evolution, structure, and function,
RL pp.110-130, Springer-Verlag, Tokyo (2000).
RN [2]
RP SEQUENCE OF 1-310 FROM N.A.
RA Wild A.;
RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: PUTATIVE ODORANT RECEPTOR.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib.ch).
CC
CC EMBL: AJ302584; CAC20504.1; -
DR EMBL: AJ302585; CAC20505.1; -
DR EMBL: AJ302586; CAC20506.1; -
DR EMBL: AJ302587; CAC20507.1; -
DR EMBL: AJ302588; CAC20508.1; -

DR EMBL: AJ302589; CAC20509.1; -
DR EMBL: AJ302590; CAC20510.1; -
DR EMBL: AJ302591; CAC20511.1; -
DR EMBL: AJ302592; CAC20512.1; -
DR EMBL: AJ302593; CAC20513.1; -
DR EMBL: 298744; CAB1427.1; -
DR InterPro: IPR00276; GPCR_Rhodopsn.
DR Pfam: PF00001; 7tm_1; 1.
DR PRINTS: PR00237; GPCR_RHODOPSIN.
DR PROSITE: PS00237; G-PROTEIN RECEPTOR FL1; 1.
DR PROSITE: PS0262; G-PROTEIN RECEPTOR FL2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein;
KW Multigene family; Olfaction; Polymorphism;
FT DOMAIN 1 25
FT TRANSMEM 26 49
FT DOMAIN 50 57
FT TRANSMEM 58 79
FT DOMAIN 80 100
FT TRANSMEM 101 120
FT TRANSMEM 121 139
FT TRANSMEM 140 158
FT DOMAIN 159 195
FT TRANSMEM 196 219
FT DOMAIN 220 236
FT TRANSMEM 237 259
FT DOMAIN 260 272
FT TRANSMEM 273 292
FT DOMAIN 293 357
FT DISULFID 97 189
FT CARBOHYD 5 5
FT VARIANT 234 234
FT
SQ SEQUENCE 357 AA; 40412 MW; 53E2F82820EC440A CRC64;
/FTID=VAR_010943.
Query Match 80.6%; Score 1295; DB 1; Length 357;
Best Local Similarity 81.3%; Pred. No. 7, 3e-90;
Matches 252; Conservative 25; Mismatches 33; Indels 0; Gaps 0;

QY 1 MNVNVDSIIIOEFLILGSDRPMLEFPLVYFLISYVTITGNTITLIVSRIDTKLHPMY 60
DB 1 MNVNVDSIIIOEFLILGSDRPMLEFPLVYFLISYVTITGNTITLIVSRIDTKLHPMY 60

QY 61 FFLTNLSLDLCYTTCTVPMOLVNLCSIRKVISYRGCAVOLFPLALGATEYLLAVMSF 120
DB 61 FFLTNLSLDLCYTTCTVPMOLVNLCSIRKVISYRGCAVOLFPLALGATEYLLAVMSF 120

QY 121 DMFVAICRPLHYSVIMHQRICLOLAASVWTGFSNSVWSTLTLOLPDCPYVIDHFCE 180
DB 121 DMFVAICRPLHYSVIMHQRICLOLAASVWTGFSNSVWSTLTLOLPDCPYVIDHFCE 180

QY 181 VPALKLISCVEETANAEELFLVSEFLPLILPLILISYATVAVARIQSAEGRQKAFGT 240
DB 181 VPALKLISCVEETANAEELFLVSEFLPLILPLILISYATVAVARIQSAEGRQKAFGT 240

QY 241 CGSHLIVSLFSTAVSVYIQPPSPSSKDGKAVSLFYGIAPMLNPLIYTLRNKEVKEG 300
DB 241 CGSHLIVSLFSTAVSVYIQPPSPSSKDGKAVSLFYGIAPMLNPLIYTLRNKEVKEG 300

QY 301 FKRIVARVFL 310
DB 301 FKRIVAKSL 310

RESULT 3
02B3_HUMAN STANDARD; PRT; 313 AA.
AC Q76000;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Olfactory receptor 2B3 (Olfactory receptor 6-4) (OR6-4) (Hs6M1-1).
GN OR2B3.


```

OS Homo sapiens (Human) .
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN (1)
RP SEQUENCE FROM N.A.
RA Evans K.;
RL Submitted (SEP-1998) to the EMBL/GenBank/DDBJ databases.
RN (2)
RP SEQUENCE FROM N.A.
RA Ziegler A., Ehlers A., Forbes S.A., Trowsdale J., Uchanska-Ziegler B.,
RA Volz A., Younger R., Beck S.;
RT "Polymorphic olfactory receptor genes and HLA loci constitute extended
haplotypes.";
RL (in) Kasanara M. (eds.);
RL Major histocompatibility complex-evolution, structure, and function,
RL pp.110-130, Springer-Verlag, Tokyo (2000).
CC -1- FUNCTION: PUTATIVE ODORANT RECEPTOR.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce
CC or send an email to license@isb.ch).
CC -----
DR EMBL; AL022273; CAAL8782.1; -
DR EMBL; AJ302537; CAC20462.1; -
DR EMBL; AJ302538; CAC20463.1; -
DR EMBL; AJ302539; CAC20464.1; -
DR EMBL; AJ302540; CAC20465.1; -
DR EMBL; AJ302541; CAC20466.1; -
DR EMBL; AJ302542; CAC20467.1; -
DR EMBL; AJ302543; CAC20468.1; -
DR EMBL; AJ302544; CAC20469.1; -
DR EMBL; AJ302545; CAC20470.1; -
DR EMBL; AJ302546; CAC20471.1; -
DR InterPro: IPR000276; GPCR_Rhodpsn.
DR Pfam: PF00001; 7tm.1; 1.
DR PROSITE; PS00237; G-PROTEIN_RECPT_F1.1; 1.
DR PROSITE; PS0262; G-PROTEIN_RECPT_F2.1; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein;
KW Multigene family; Olfaction.
FT DOMAIN 1 25 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 26 49 1 (POTENTIAL).
FT DOMAIN 50 57 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 58 79 2 (POTENTIAL).
FT DOMAIN 80 100 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 101 120 3 (POTENTIAL).
FT DOMAIN 121 139 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 140 158 4 (POTENTIAL).
FT DOMAIN 159 195 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 196 219 5 (POTENTIAL).
FT DOMAIN 220 236 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 237 259 6 (POTENTIAL).
FT DOMAIN 260 272 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 273 292 7 (POTENTIAL).
FT DOMAIN 293 313 CYTOPLASMIC (POTENTIAL).
FT DISULFD 97 189 BY SIMILARITY.
FT CARBOHD 5 5 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT SEQUENCE 313 AA; 35542 MW; 768BF1EB/B624440 CRC64;

Query Match 70.3%, Score 1130; DB 1; Length 313;
Best Local Similarity 70.3%; Pred. No. 1.3e-77;
Matches 220; Conservative 34; Mismatches 59; Indels 0; Gaps 0;

07 1 MNMVDSTIQEIIILGFSRPMLEPLLVPLISTVTVIFGMLITILYSRLDTKATPMY 60
1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1 MNMVESSPKREIILGFSRPMLEPLLVPLISTVTVIFGVSIMMVCIDPKKATPMY 60

```

QY	61	FEFLNLSLDDLCYTCYTPOMLVNCSIRKIVSYSGVAAOEIFALGATEYLLAAVNSF	120
QY	61	FEFLNLSLDDLCYTCYTPOMLVNCSIRKIVSYSGVAAOEIFALGATEYLLAAVNSF	120
Db	61	FEFLNLSLDDLCYTCYTTTYPHMLVNIQCCKKTTSTAGCAVHAHLIFALATECILLAAVNSF	120
QY	121	DMFVAICRPLRHSYVMHQRCLQLAAASWYTGFSNSVM,STLTQLDPLCDPYVDHFICE	180
Db	121	DMFVAICRPLRHSYVMHQRCLQLAAASWYTGFSNSVM,STLTQLDPLCDPYVDHFICE	180
QY	161	VPALLKLSCEVETANEAELFVSELFHLIPLTLIIISAFYRAVLRLQSAEGRKARAGT	240
Db	161	VPALLKLSCEVETANEAELFVSELFHLIPLTLIIISAFYRAVLRLQSAEGRKARAGT	240
QY	241	CGSHLIIVVSLFSTPVSAYLDPSPSSKDOGMMVSLFGIITAPMNPILYTLRNVEVGE	300
Db	241	CGSHLIIVVSLFSTPVSAYLDPSPSSKDOGMMVSLFGIITAPMNPILYTLRNVEVGE	300
QY	301	FKRLVARVELIRK 313	
Db	301	FKRLMPRIEFCK 313	
RESULT	4		
OC	02J3_HUMAN	STANDARD;	PRT; 311 AA.
AC	076001; Q9GZK5; Q9GZL4; Q9GZL5;		
DC	16-OCT-2001 (Rel. 40, Created)		
DT	16-OCT-2001 (Rel. 40, Last sequence update)		
DT	16-OCT-2001 (Rel. 40, Last annotation update)		
DE	Olfactory receptor 2J3 (Olfactory receptor 6-6) (OR6-6) (Hs6M1-3).		
GN	OR2J3.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
CC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
OX	NCBI_TaxID=9606;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RA	Evans K.;		
RL	Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.		
RN	[2]		
RA	SEQUENCE FROM N.A., AND VARIANTS A-113; Q-226; T-228 AND M-261.		
RA	Ziegler A., Ehlers A., Forbes S.A., Trowsdale J., Uchanska-Ziegler B.,		
RA	Volz A., Younger R., Beck S.;		
RT	"Polymorphic olfactory receptor genes and HLA loci constitute extended		
RT	haplotypes.";		
RL	(In) Kasahara M. (eds.);		
RL	Major histocompatibility complex-evolution, structure, and function,		
RL	pp.110-130, Springer-Verlag, Tokyo (2000).		
CC	-1- FUNCTION: PUTATIVE ODORANT RECEPTOR.		
CC	-1- SUBCELLULAR LOCATION: Integral membrane protein.		
CC	-1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration		
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -		
CC	the European Bioinformatics Institute. There are no restrictions on its		
CC	use by non-profit institutions as long as its content is in no way		
CC	modified and this statement is not removed. Usage by and for commercial		
CC	entities requires a license agreement (see http://www.isb-sib.ch/announce/		
CC	or send an email to license@isb-sib.ch).		
CC	-----		
DR	EMBL; AL022727; CAA18783.1;		
DR	EMBL; AJ302547; CAC20472.1;		
DR	EMBL; AJ302548; CAC20473.1;		
DR	EMBL; AJ302549; CAC20474.1;		
DR	EMBL; AJ302550; CAC20475.1;		
DR	EMBL; AJ302551; CAC20476.1;		
DR	EMBL; AJ302552; CAC21440.1;		
DR	EMBL; AJ302553; CAC21441.1;		
DR	EMBL; AJ302554; CAC21442.1;		
DR	EMBL; AJ302555; CAC21443.1;		
DR	EMBL; AJ302556; CAC21444.1;		
DR	EMBL; AJ302557; CAC20477.1;		
DR	EMBL; AJ302558; CAC20478.1;		

DR InterPro: IPR000276; GPCR_Rhodpsn.
 DR Pfam: PF00001; 7tm.1; 1.
 DR PROSITE: PS00237; G_PROTEIN_RECIP_F1_1; 1.
 DR PROSITE: PS0262; G_PROTEIN_RECIP_F1_2; 1.
 DR G-protein coupled receptor; Transmembrane; Glycoprotein;
 KW Multigene family; Olfaction; Polymorphism.
 FT DOMAIN 1 28
 FT TRANSMEM 29 52
 FT DOMAIN 53 60
 FT TRANSMEM 61 82
 FT DOMAIN 83 103
 FT TRANSMEM 104 123
 FT DOMAIN 124 142
 FT TRANSMEM 143 161
 FT DOMAIN 162 198
 FT TRANSMEM 199 222
 FT DOMAIN 223 239
 FT TRANSMEM 240 262
 FT DOMAIN 263 275
 FT TRANSMEM 276 295
 FT DOMAIN 296 311
 FT DISULFID 100 192
 FT CARBOHYD 8 8
 FT VARIANT 113 113
 FT VARIANT 226 226
 FT VARIANT 228 228
 FT VARIANT 261 261
 FT SEQUENCE 311 AA: 34950 MW: 842011459955A68 CRC64;
 Query Match 58.1%; Score 933; DB 1; Length 311;
 Best Local Similarity 59.9%; Pred. No. 6.3e-63;
 Matches 181; Conservative 43; Mismatches 78; Indels 0; Gaps 0;

QY 4 VNSDSTIOEILGFSDRPVLEFLLVFLISTVTITFGNLTIIIVSRDLTKHTPWYFL 63
 DB 7 VASSEGEYVILVGFNSWPLLEVVIFVVLITMLTIGLFIITLISYDLSHTPTWYFL 66
 QY 64 TNLSDLCYTTCTVPMQVNLVNCSTIKVISTYGCVAQLEFIALGTERELLVAWSDFW 123
 DB 67 SNLSRELDYTTSSIPOLVNLNMGPEKTSYAGCMQLTFVLAIGTTECVLLVWMSYDRY 126
 QY 124 VVPCPLHYSVTHORLCIQLAAASWVTGFSNVMSTLTLDLPIDDPVIDHIFCEVPA 183
 DB 127 AAVCPRLHVTVMHPRFCHLLAVASWVSGFTNSALHSTTFWVPLCGHQQVDHFECEVPA 186
 QY 184 LKLSCEVTANAEALFLVSELFHLIPLTLILISYAFIVRAVLRIQSAEGROKAFEGTGS 243
 DB 187 LRLSCVDTHVNEFLMTLITSSIFVLIPILTLISYGAIVRAVLRMQSTGLQKVEGTCA 246
 QY 244 HLIVVSLPSTAVSYVLDPPSPSSKDGKMSVLFGLIAPMLNPLTYLTRNKEVKEGFR 303
 DB 247 HLMAVSLFFIPAMCILOPPSGNSQDGKFIALFTVVPISNPLTYLTRNKEVKEGFR 306
 QY 304 LV 305
 DB 307 LM 308

RESULT 5
 02J2_HUMAN STANDARD: PRT: 312 AA.
 AC 076002; Q9GZK2; Q9GZL3;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DE 16-OCT-2001 (Rel. 40, Last annotation update)
 GN Olfactory receptor 2J2 (Olfactory receptor 6-8) (OR6-8) (Hs6ml-6).
 OS Homo sapiens (Human).

CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Evans K.;
 RN Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
 RP [2]
 RA Ziegler A., Ehlers A., Forbes S.A., Trowsdale J., Uchanska-Ziegler B.,
 RA Vozl A., Younger R., Beck S.;
 RT "Polymorphic olfactory receptor genes and HLA loci constitute extended
 haplotypes.";
 RL (in) Kasahara M. (eds.);
 RL Major histocompatibility complex-evolution, structure, and function,
 RL pp.110-130, Springer-Verlag, Tokyo (2000).
 CC -1- FUNCTION: PUTATIVE ODORANT RECEPTOR.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: AL022727; CAA18784.1; -
 DR EMBL: AJ302571; CAC20491.1; -
 DR EMBL: AJ302572; CAC20492.1; -
 DR EMBL: AJ302573; CAC20493.1; -
 DR EMBL: AJ302574; CAC20494.1; -
 DR EMBL: AJ302575; CAC20495.1; -
 DR EMBL: AJ302576; CAC20496.1; -
 DR EMBL: AJ302577; CAC20497.1; -
 DR EMBL: AJ302578; CAC20498.1; -
 DR EMBL: AJ302579; CAC20499.1; -
 DR EMBL: AJ302580; CAC20500.1; -
 DR EMBL: AJ302581; CAC20501.1; -
 DR EMBL: AJ302582; CAC20502.1; -
 DR EMBL: AJ302583; CAC20503.1; -
 DR InterPro: IPR000276; GPCR_Rhodpsn.
 DR Pfam: PF00001; 7tm.1; 1.
 DR PROSITE: PS00237; G_PROTEIN_RECIP_F1_1; 1.
 DR PROSITE: PS0262; G_PROTEIN_RECIP_F1_2; 1.
 DR G-protein coupled receptor; Transmembrane; Glycoprotein;
 KW Multigene family; Olfaction; Polymorphism.
 FT DOMAIN 1 26
 FT TRANSMEM 27 50
 FT DOMAIN 51 58
 FT TRANSMEM 59 80
 FT DOMAIN 81 101
 FT TRANSMEM 102 121
 FT DOMAIN 122 140
 FT TRANSMEM 141 159
 FT DOMAIN 160 196
 FT TRANSMEM 197 220
 FT DOMAIN 221 237
 FT TRANSMEM 238 260
 FT DOMAIN 261 273
 FT TRANSMEM 274 293
 FT DISULFID 294 312
 FT CARBOHYD 98 190
 FT VARIANT 6 6
 FT VARIANT 74 74
 FT VARIANT 111 111
 FT VARIANT 146 146
 FT VARIANT 218 218

/FTID-VAR_010945.
 A -> T (IN 6M1-6*03).
 /FTID-VAR_010946.
 A -> V (IN 6M1-6*02 AND 6M1-6*03).
 /FTID-VAR_010947.
 A -> T (IN 6M1-6*02 AND 6M1-6*03).
 /FTID-VAR_010948.

SQ SEQUENCE 312 AA; 35204 MW; FB7D20BB2379C43E CRC64;
 Query Match 56.3%; Score 904; DB 1; Length 312;
 Best Local Similarity 59.9%; Pred. No. 9.2e-61;
 Matches 176; Conservative 43; Mismatches 75; Indels 0; Gaps 0;

QY 12 FILFGSDPRWLEFPLLVFLISYVTITFGNTLIIVSRDLTKLHPMPFELTNLSLDL 71
 DB 13 FILFGSDPRWLEFPLLVFLISYVTITFGNTLIIVSRDLTKLHPMPFELTNLSLDL 72
 QY 72 CYTTCVPMQMLVNCISIRKIVSYRGCAVOLFIFALGATEYLLAAVSPFMAICRPLH 131
 DB 73 CRTTSSIPDLVNLKRGPEKTIISAGCMVQLYEVLALGACVLLVMSYDRYVAVCRPLH 132
 QY 132 YSVIMHORLCLOLAAASWVTGFSNSVWLSLTLLQLPLCDPYVIDHFLCEVPALLKISCVE 191
 DB 133 YVIMHPRCHLLAAASWVIGFTISALHSSFTFWVPLCGHRLVDHFCFEPALLRISCVD 192
 QY 192 TTANAELEFVSEFLPLPLTLILISYAFIVRAVLRIGSAEGKAKFGTCGSHLYVSLF 251
 DB 193 TTANAELEFVSEFLPLPLTLILISYAFIVRAVLRIGSAEGKAKFGTCGSHLYVSLF 252
 QY 252 YSNAVSYLQPPSPSSKDGKMWLSLEFGIAPMLNPLIYLRKKEVKEGKRLV 305
 DB 253 YSNAVSYLQPPSPSPSSKDGKMWLSLEFGIAPMLNPLIYLRKKEVKEGKRLV 306

RESULT 6
 OL15_MOUSE STANDARD; PRT; 312 AA.
 AC P23275;
 DT 01-MAR-1992 (Rel. 21, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE Olfactory receptor 15 (OR3).
 GN OLFR15.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxId=10090;
 RN [1]
 RP MEDLINE-93028384; PubMed-1384038;
 RA Neft P., Hermans-Borgmeyer I., Artieres-pin H., Beasley L.,
 RA Dionne V.E., Heinemann S.F.;
 RT "Spatial pattern of receptor expression in the olfactory epithelium.";
 RT Proc. Natl. Acad. Sci. U.S.A. 89:8948-8952(1992).
 CC -1- FUNCTION: PUTATIVE ODORANT RECEPTOR.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- TISSUE SPECIFICITY: OLFACTORY EPITHELIUM.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M84005; AAA39862.1; -;
 DR GCRDB; GCR_0494; -;
 DR MGD; MGI:106182; Olfr15.
 DR InterPro; IPR000276; GPCR_Rhodpsn.
 DR Pfam; PF00001; 7tm1.1;
 DR PRINTS; PR00245; OLFACTORYR.
 DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1;
 DR PROSITE; PS0262; G_PROTEIN_RECEP_F1_2; 1;
 DR G-protein coupled receptor; Transmembrane; Glycoprotein;
 KM Olfaction; Multigene family.
 FT DOMAIN 1 24
 FT TRANSMEM 25 48
 FT TRANSMEM 1 (POTENTIAL).

FT DOMAIN 49 57
 FT TRANSMEM 58 79
 FT DOMAIN 80 100
 FT TRANSMEM 101 120
 FT DOMAIN 121 139
 FT TRANSMEM 140 160
 FT DOMAIN 161 200
 FT TRANSMEM 201 222
 FT DOMAIN 223 236
 FT TRANSMEM 237 261
 FT DOMAIN 273 292
 FT TRANSMEM 293 312
 FT DISULFID 97 189
 FT CARBOHYD 6 6
 SQ SEQUENCE 312 AA; 34333 MW; 8D3877EBB51E132 CRC64;

Query Match 56.1%; Score 902; DB 1; Length 312;
 Best Local Similarity 58.8%; Pred. No. 1.3e-60;
 Matches 174; Conservative 46; Mismatches 76; Indels 0; Gaps 0;

QY 12 FILFGSDPRWLEFPLLVFLISYVTITFGNTLIIVSRDLTKLHPMPFELTNLSLDL 71
 DB 12 FILFGSDPRWLEFPLLVFLISYVTITFGNTLIIVSRDLTKLHPMPFELTNLSLDL 71
 QY 72 CYTTCVPMQMLVNCISIRKIVSYRGCAVOLFIFALGATEYLLAAVSPFMAICRPLH 131
 DB 72 ARTTSSVPMKRLKMLNGPDKTISYGGCYTQLYEFLMLGATCILLVMAFPRYVAVCRPLH 131
 QY 132 YSVIMHORLCLOLAAASWVTGFSNSVWLSLTLLQLPLCDPYVIDHFLCEVPALLKISCVE 191
 DB 132 YVIMHPRCHLLAAASWVIGFTISALHSSFTFWVPLCGHRLVDHFCFEPALLRISCVD 191
 QY 192 TTANAELEFVSEFLPLPLTLILISYAFIVRAVLRIGSAEGKAKFGTCGSHLYVSLF 251
 DB 192 TTANAELEFVSEFLPLPLTLILISYAFIVRAVLRIGSAEGKAKFGTCGSHLYVSLF 251
 QY 252 YSNAVSYLQPPSPSSKDGKMWLSLEFGIAPMLNPLIYLRKKEVKEGKRLV 307
 DB 252 YSNAVSYLQPPSPSSKDGKMWLSLEFGIAPMLNPLIYLRKKEVKEGKRLV 307

RESULT 7
 OL2H2_HUMAN STANDARD; PRT; 312 AA.
 ID 02H2_HUMAN
 AC 095018;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Olfactory receptor 2H2 (Hs6M1-12).
 GN OR2H2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxId=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Younger R.;
 RL Submitted (JAN-1999) to the EMBL/Genbank/DBJ databases.
 CC -1- FUNCTION: PUTATIVE ODORANT RECEPTOR.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AL031983; CAA21455.1; -;

DR InterPro: IPR000276; GPCR_Rhodopsn.
 DR Pfam: PF00001; 7tm_1; 1.
 DR PRINTS: PR00237; GPCR_RHODOPSIN.
 DR PROSITE: PS00237; G_PROTEIN_REC_P1.1; 1.
 DR PROSITE: PS00262; G_PROTEIN_REC_P1.2; 1.
 DR G-protein coupled receptor; Transmembrane; Glycoprotein;
 KM Multigene family; Olfaction.
 FT DOMAIN 1 23 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 24 47 1 (POTENTIAL).
 FT DOMAIN 48 55 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 56 77 2 (POTENTIAL).
 FT DOMAIN 78 98 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 99 118 3 (POTENTIAL).
 FT DOMAIN 119 137 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 138 156 4 (POTENTIAL).
 FT DOMAIN 157 193 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 194 217 5 (POTENTIAL).
 FT DOMAIN 218 234 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 235 257 6 (POTENTIAL).
 FT DOMAIN 258 270 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 271 290 7 (POTENTIAL).
 FT DOMAIN 291 312 CYTOPLASMIC (POTENTIAL).
 FT DISULFID 95 187 BY SIMILARITY.
 FT CARBOHYD 3 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT SEQUENCE 312 AA; 34797 MW; 2C7AB6C7E08AB6B CRC64;

Query Match 55.3%; Score 889; DB 1; Length 312;
 Best Local Similarity 56.2%; Pred. No. 1.2e-59;
 Matches 171; Conservative 50; Mismatches 83; Indels 0; Gaps 0;

QY 4 VNDSTIQTETLLGSDRPMLEPPLVFLISTVTGNTITLVRDLTKLTMPYFEL 63
 DB 2 VNSSTPGFLLLGFSHPGLERTLFVVFSTYLLTLVAGLTLALPKLHSPMYFEL 61
 QY 64 TNSLTDLCYTTCTVQPMVLNCSIRKVSIRGCVAGLFFPLAGREYLLAVMSDFW 123
 DB 62 SNISFLDCTTSCVPMNLNMGPKTTSFLDCSVQITFLSLGTETCLLTVMADRY 121
 QY 124 VAICRPLHYSVIMHQRILQLOLAASVWTFGSNSWMLSTLLQLPLDDPYVDHFLCEVPA 183
 DB 122 VAVCGPLHATITIHRLCQGLASVAVWIGLVESVQTPSLHLPFCPDQGVDFVCEVPA 181
 QY 184 LKLSCEVTANAEFLVSEFLHPLILISYAFYRAVRIOSAGRCQAFGTCS 243
 DB 182 LRLSEDSYSEIDQVAVSVFLVPLSLIVSYGAIWAVYRINSAGRRKAFGTCS 241
 QY 244 HLIIVSLFSTAVSYVLOPPSPSSKDGKWSLFGYLIAPMLPLIYTLRNKVKSGEKR 303
 DB 242 HLIIVTLFTSVIAYVLOPKNPYADRGKFFGLFVAGTSPSLPLIYTLRNKEVTRAFRR 301
 QY 304 LVAR 307
 DB 302 LGK 305

RESULT 8
 O2H3_HUMAN STANDARD; PRT; 316 AA.
 AC 015062;
 DT 01-NOV-1997 (Rel. 35; Created)
 DT 16-OCT-2001 (Rel. 40; Last sequence update)
 DT 16-OCT-2001 (Rel. 40; Last annotation update)
 DE Olfactory receptor 2H3 (Olfactory receptor-like protein FAT11).
 GN OR2H3 OR OLFR2 OR FAT11.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95394447; PubMed=7665158;
 RA Fan W., Liu Y.-C., Parimoo S., Weissman S.M.;

RT "Olfactory receptor-like genes are located in the human major
 RT histocompatibility complex.";
 RL genomics 27:119-123(1995).
 RN [2]
 RP SEQUENCE OF 10-312 FROM N.A. AND VARIANTS S-30; V-48 AND A-220.
 RX MEDLINE=20341390; PubMed=10860742;
 RA Eklund A.C., Belchak M.M., Lapidus K., Raha-Chowdhury R., Ober C.;
 RT "Polymorphisms in the HLA-linked olfactory receptor genes in the
 RT Hutterites.";
 RT Hum. Immunol. 61:711-717(2000).
 RL -1- FUNCTION: PUTATIVE ODORANT RECEPTOR.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: L35475; AAB36567.1; -;
 DR EMBL: AF211939; AAF98751.1; -;
 DR EMBL: AF211940; AAF98752.1; -;
 DR EMBL: AF211941; AAF98753.1; -;
 DR EMBL: AF211942; AAF98754.1; -;
 DR GCRDB: GCR_1814; -;
 DR MIM: 600578; -;
 DR InterPro: IPR000276; GPCR_Rhodopsn.
 DR Pfam: PR00237; GPCR_RHODOPSIN.
 DR PRINTS: PR00237; GPCR_RHODOPSIN.
 DR PROSITE: PS00237; G_PROTEIN_REC_P1.1; 1.
 DR PROSITE: PS00262; G_PROTEIN_REC_P1.2; 1.
 KM G-protein coupled receptor; Transmembrane; Glycoprotein;
 KM Multigene family; Olfaction.
 FT DOMAIN 1 22 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 23 46 1 (POTENTIAL).
 FT DOMAIN 47 55 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 56 77 2 (POTENTIAL).
 FT DOMAIN 78 98 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 99 118 3 (POTENTIAL).
 FT DOMAIN 119 137 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 138 158 4 (POTENTIAL).
 FT DOMAIN 159 198 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 199 220 5 (POTENTIAL).
 FT DOMAIN 221 234 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 235 259 6 (POTENTIAL).
 FT DOMAIN 260 270 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 271 290 7 (POTENTIAL).
 FT DOMAIN 291 316 CYTOPLASMIC (POTENTIAL).
 FT DISULFID 95 187 BY SIMILARITY.
 FT CARBOHYD 3 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT VARIANT 30 30 F -> S.
 FT VARIANT 48 48 A -> V.
 FT VARIANT 220 220 /FTID=VAR_010228.
 FT VARIANT 220 220 T -> A.
 FT /FTID=VAR_010229.
 FT E -> G (IN REF. 1).
 FT V -> D (IN REF. 1).
 FT F -> I (IN REF. 1).
 FT P -> T (IN REF. 1).
 FT L -> A (IN REF. 1).
 FT Lf -> MK (IN REF. 1).
 FT E -> G (IN REF. 1).
 FT KGR -> TWM (IN REF. 1).
 FT E -> G (IN REF. 1).
 FT I -> V (IN REF. 1).
 FT VT -> IK (IN REF. 1).
 FT F -> L (IN REF. 1).
 FT ROSRE -> MGITO (IN REF. 2).
 FT SEQUENCE 316 AA; 35410 MW; BE29B672B663ATCE CRC64;

```

Query Match          55.1% Score 885; DB 1; Length 316;
Best Local Similarity 56.1%; Pred. No. 2,4e-59;
Matches 170; Conservative 50; Mismatches 83; Indels 0; Gaps 0;

OY      5 NDSIIQEFILLGFSDRPMLEFPLLVFLVFIISYTVTFEGNMTIIVSRIDTKLHTPMPFYLT 64
           |||::|||::| | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB       3 NQSSPPGFLLLGFSGSHPELRTLPVVFTSYLLTLVGNTLIILIALADPKHSPKTYEELS 62
           |||::|||::| | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY      65 NLSLDLCYTCYPQMVLVNLCSIRKVISYGCVAQLFITFALGATEYLIAVMSDFWV 124
           |||::|||::| | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB       63 NLSEFDLCFTTSCVPQMLVNLMGPCKRTISFDICSVQIPIFLSTGTCECILITVMAFDXYV 122
           |||::|||::| | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY      125 AICRFHSYSVMIOHOLACGLAAASVWFNSGVMSWTSTLIDPCDDPVVIDHFICEVAL 184
           |||::|||::| | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB       123 AVCGPLHATITIIHPRLCQMLASVANVICLVESVOPTBSTLHPCPDPQVDVCEVFAL 182
           |||::|||::| | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY      185 LKLSCETTANAEELFVSELFEHLIPLLILISAFIVRAVLRIOSAGROKAFGTGGSH 244
           |||::|||::| | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB       183 IRLSCEPTSYNEIQVAVASVFILVPLSLILVSYGATIMAVLRINSAGRKRAFGTCCSH 242
           |||::|||::| | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY      245 LIVLSFSTAVSYVLQPPSPSSDKDGKMSLFPGITAPMLNPPLYTLRNKEVKEGFRRL 304
           |||::|||::| | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB       243 LVVLFLEFSVIAYVLQRPKNPADEGRKFGLFYAVGTPSNPLLYTLRNKEVTRAFRRLL 302
           |||::|||::| | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY      305 VAR 307
           : :
DB       303 LGK 305

RESULT    9
OZC1_HUMAN OZC1_HUMAN STANDARD; PRT; 312 AA.
AC         095371;
DT      16-OCT-2001 (Rel. 40, Created)
DT      16-OCT-2001 (Rel. 40, Last sequence update)
DT      16-OCT-2001 (Rel. 40, Last annotation update)
DE      Olfactory receptor 2C1 (OLFml3).
GN      OR2C1.
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX      NCBI_TaxId=9606;
RX      [1]
RX      SEQUENCE FROM N.A.
RX      MEDLINE=98063798; PubMed=9847080;
RA      Centola M., Chen X., Sood R., Deng Z., Akseiljevich I., Blake T.,
RA      Rieke D., Chen X., Wood G., Zaks N., Richards N., Krizman D.,
RA      Mansfield E., Apostolou S., Liu J., Shatran N., Vedula A., Hamon M.,
RA      Cereek A., Khan T., Gumucio D., Callen D.P., Richards R.I.,
RA      Mojzis R.K., Doggett N.A., Collins F.S., Liu P.P.,
RA      Fischel-Ghodasian N., Kastner D.L.;
RT      "Construction of an approximately 700 kb transcript map around the
RL      familial mediterranean fever locus on human chromosome 16p13.3.";
RL      Genome Res. 9:1-21(1998).
CC      -|- FUNCTION: PUTATIVE ODORANT RECEPTOR.
CC      -|- SUBCELLULAR LOCATION: Integral membrane protein.
CC      -|- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outpost. -|-
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL: AF098664; AAC83557.1; -
DR      InterPro: IPR000276; GPCR_Rhodopsn.
DR      Pfam: PF00001; 7tm_1; 1.
DR      PRINTS: PR00237; GPCRHHODOPSN.
DR      PROSITE: PS00237; G_PROTEIN_RECPT_FL_1; 1.
```

DR	PROSITE: PS0262; G-PROTEIN RECEPTOR FL2; 1.	
KW	G-protein coupled receptor; Transmembrane; Glycoprotein;	
KV	MultiGene family; Olfaction.	
FT	DOMAIN 1 25	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM 26 49	1 (POTENTIAL).
FT	DOMAIN 50 57	CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM 58 79	2 (POTENTIAL).
FT	DOMAIN 80 100	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM 101 120	3 (POTENTIAL).
FT	DOMAIN 121 139	CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM 140 158	4 (POTENTIAL).
FT	DOMAIN 159 196	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM 197 219	5 (POTENTIAL).
FT	DOMAIN 220 236	CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM 237 259	6 (POTENTIAL).
FT	DOMAIN 260 272	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM 273 292	7 (POTENTIAL).
FT	DOMAIN 293 312	CYTOPLASMIC (POTENTIAL).
FT	DISULFD 97 189	BY SIMILARITY.
FT	CARBOHD 5	N-LINKED (GLCNAC. .) (POTENTIAL).
SO	SEQUENCE 312 AA: 34506 MW: CCADYD349BEB9E7 CR664;	

```

Query Match          54.9%:  Score 883;  DB 1:  Length 312;
Match Local Similarity 55.7%:  Pred. No. 3,4e-59;
Matches 171;  Conservative 51;  Mismatches 85;  Indels 0;  Gaps 0

OY      1  MNMVDSIIQIEFILLGFSDRPWMEFPLVVFVIFSYVTFNGMLITIIIVSRIDPTKLTTPMY 60
      1  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      1  MGVVDSLSIQGVIVLMSISDHPLEMFIFAILFVSYTIIILGTSIIILSRLEARLHTTPY 60
      1  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

OY      61  PFLTMSLIDLCYTTCTYPOMLVNLCSIRKVIYSRCGVAOLFIFLALGATEFLLAAVMSF 120
      61  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      61  PFLSMISSDILFAFATTSVPQMLINMGPEKTIISYGGCIQTLVFLMFGATECILLVMAF 120
      61  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

OY      121  DFEVACIRPHYSVIMHOLICQLAAASVNGEFSNVMSTYTLLOPLCDPVIDHFLCE 180
      121  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      121  DRVAVACRPLRTAIINPOLCMLLAVIAMLGIGLSVYSTTTLLOPLDGHRRVSEFLCE 180
      121  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

OY      181  VPALKILSCVETTANEAELFVSEFLHILPLTLLISYAVIRAVLRIOASAGROKAFGT 240
      181  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      181  VPAMIKILCGDSTLNOAVLNGVCEFTTAVPLISYIVCYLLIAQVAVLKISAGRRKAFMT 240
      181  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

OY      241  CGSHLIIVSLFSTAVSVYLLQPPSPSSKQGGKMSVLFYIIMPLNPLTYTLRNKEVKS 300
      241  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      241  CISHLIIVLFLEFGSASYGLLPAKNSKODQGFISLFSVLYPMVNPPLTYTLRNNEVKA 300
      241  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

OY      301  FKRLVAR 307
      301  : : : : :
Db      301  LRRLLGK 307
      301  : : : : :

RESULT 10
O2W1_HUMAN
AC      O2W1_HUMAN          STANDARD:          PRT:          320 AA.
ID      O9Y3N9: 09GZLL:
DT      16-OCT-2001 (Rel. 40, Created)
DT      16-OCT-2001 (Rel. 40, Last sequence update)
DT      16-OCT-2001 (Rel. 40, Last annotation update)
DE      Olfactory receptor 2W1 (Hs6M1-15).
GN      OR2W1.
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX      NCBI_TaxID=9606;
RN      [1]
RP      SEQUENCE FROM N.A.
RA      Smith S.;
RL      Submitted (MAY-1999) to the EMBL/Genbank/DBJ databases.
[2]
RP      SEQUENCE FROM N.A., AND VARIANTS VAL-81 AND ASN-236.
RA      Ziegler A., Ehlers A., Forbes S.A., Trowsdale J., Uchanska-Ziegler B.,
RA      Volz A., Younger R., Beck S.;

```

RT "Polymorphic olfactory receptor genes and HLA loci constitute extended haplotypes." ;

RL (in) Kasahara M. (eds.);

RL Major histocompatibility complex-evolution, structure, and function, pp.110-130, Springer-Verlag, Tokyo (2000).

CC -1- FUNCTION: PUTATIVE ODORANT RECEPTOR.

CC -1- SUBCELLULAR LOCATION: Integral membrane protein.

CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC

DR EMBL: AJ035402; CAB42853.1; -

DR EMBL: AJ302594; CAC20514.1; -

DR EMBL: AJ302595; CAC20515.1; -

DR EMBL: AJ302596; CAC20516.1; -

DR EMBL: AJ302597; CAC20517.1; -

DR EMBL: AJ302598; CAC20518.1; -

DR EMBL: AJ302599; CAC20519.1; -

DR EMBL: AJ302600; CAC20520.1; -

DR EMBL: AJ302601; CAC20521.1; -

DR EMBL: AJ302602; CAC20522.1; -

DR EMBL: AJ302603; CAC20523.1; -

DR InterPro: IPR000276; GPCR_Rhodopsin.

DR Pfam: PF00001; 7tm.1.1.

DR PRINTS: PR00237; GPCR_Rhodopsin.

DR PROSITE: PS00237; G-PROTEIN_RECPT_FL_1; 1.

DR PROSITE: PS0262; G-PROTEIN_RECPT_FL_2; 1.

KM G-protein coupled receptor: Transmembrane; Glycoprotein;

KM Multigene family; Olfaction; Polymorphism.

FT DOMAIN 1 25

FT TRANSMEM 26 49

FT DOMAIN 50 57

FT TRANSMEM 58 79

FT DOMAIN 80 100

FT TRANSMEM 101 120

FT DOMAIN 121 139

FT TRANSMEM 140 158

FT DOMAIN 159 195

FT TRANSMEM 196 219

FT DOMAIN 220 236

FT TRANSMEM 237 259

FT DOMAIN 260 272

FT TRANSMEM 273 292

FT DOMAIN 293 320

FT DISULFID 97 189

FT CARBOHYD 5 5

FT VARIANT 81 81

FT VARIANT 296 296

FT SEQUENCE 320 AA; 36101 MW; 7ABECD3F443E1D0 CRC64;

QY Query Match 54.2%; Score 871; DB 1; Length 320;

QY Best Local Similarity 55.1%; Pred. No. 2.7e-58;

QY Matches 168; Conservative 51; Mismatches 86; Indels 0; Gaps 0;

DB 1 MNVNDIIIOEFLILGSDRPMLEPFLVLFVLTIRGNLTITIVSLDKLHPMY 60

DB 1 MDSNYSLSLHGFLLGFSNFKHEMLISGVAFYLLTLLGNATLILASLDSQLHPMY 60

QY 61 FFLTNLSLDLCYCTCTPQMLVNLCSIRKVISYRCVADLFELALGATEYLLAVMSF 120

DB 61 FFLRNLSFLDLCFTTSTIIPQMLVNLGMPDKTISYVGGIIDLVMYMGVSCLLAVMSY 120

QY 121 DMFVALCRPLHYVIMHQRCLDLAASVWTGSSNWSWLTLLQLPLCPYVIDHFLCE 180

DB 121 DRETAICKPLHYEVMNPHCLKMIIMWISLANSVVLCTLLNLPCCGNNIIDHFLCE 180

QY 181 VPALLKLSCEVETPANEAELEFLVSELFLPILPTLLISYAFVRAVLRIQSAEGROKAFGT 240

DB 181 LPAVLKACADDTTIVEMSVFALGIILVPLILLLISGYIAKAVLTKSAASROKAMNT 240

QY 241 CGSHLIVSVSFYSTAVSVYIQPPSSSKDQKWSLEFYGIAPMLNPLIYLRNKEVKEG 300

DB 241 CGSHLIVSVMSFEYGIIVMYIQGNRASKDQKFLTEFYVTPSLNPLIYLRNKMKA 300

QY 301 FKRLV 305

DB 301 LKKLM 305

RESULT 11

02HL_HUMAN

ID 02HL_HUMAN STANDARD; PRT; 316 AA.

AC Q9GZK4; Q9GZK9; O43661; O43629;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 01-MAR-2002 (Rel. 41, Last annotation update)

DE Olfactory receptor 2H1 (Hs6M1-16) (Olfactory receptor 6-2) (OR6-2)

DE (OLFR42A-9004.14/9026.2).

GN OR2HL.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

OX NCBI_TaxId=9606;

RN [1]

RP SEQUENCE FROM N.A., AND VARIANT ASN-63.

RA Ziegler A., Ehlers A., Forbes S.A., Trowsdale J., Uchanska-Ziegler B., Volz A., Younger R., Beck S.;

RT Polymorphic olfactory receptor genes and HLA loci constitute extended haplotypes." ;

RL (in) Kasahara M. (eds.);

RL Major histocompatibility complex-evolution, structure, and function, pp.110-130, Springer-Verlag, Tokyo (2000).

RN [2]

RP SEQUENCE FROM N.A.

RA Whitaker H.;

RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.

RN [3]

RP SEQUENCE OF 58-251 FROM N.A.

RA Galliano H.;

RT "Olfactory receptor gene cluster in man and mouse major histocompatibility complex." ;

RL Immunogenetics 0:0-0(1998).

RN [4]

RP SEQUENCE OF 58-251 FROM N.A.

RA Amadou C., Avoustin P., Ribouchon M.T., Bouissou C., Tazi Ahnini R., Ayer C., Portarotti P.;

RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.

CC -1- FUNCTION: PUTATIVE ODORANT RECEPTOR.

CC -1- SUBCELLULAR LOCATION: Integral membrane protein.

CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC

DR EMBL: AJ302604; CAC20524.1; -

DR EMBL: AJ302605; CAC20525.1; -

DR EMBL: AJ302606; CAC20526.1; -

DR EMBL: AJ302607; CAC20527.1; -

DR EMBL: AJ302608; CAC20528.1; -

DR EMBL: AJ302609; CAC20529.1; -

DR EMBL: AJ302610; CAC20530.1; -

DR EMBL: AJ302611; CAC20531.1; -

DR	EMBL:	AJ302612;	CAC20532.1;	-
DR	EMBL:	AJ302613;	CAC20533.1;	-
DR	EMBL:	AL035542;	CAB44506.1;	-
DR	EMBL:	AF040078;	AAC00184.1;	-
DR	EMBL:	AF044491;	AAC00188.1;	-
DR	InterPro:	IPR000276;	GPCRRhodopsin.	
DR	Pfam:	PF00001;	7tm_1.	1.
DR	PRINTS:	PR00337;	GPCRRHODOPSIN.	
DR	PROSITE:	PS00337;	G_PROTEIN_RECEP_F1_1;	1.
DR	PROSITE:	PS00622;	G_PROTEIN_RECEP_F1_2;	1.
KW	G-protein coupled receptor; Transmembrane; Glycoprotein Multigene family; Olfaction; Polymorphism. D -> N (IN 6M1-16*02).			
KW	VARIANT	63	63	/FOLD-VAR.010944.
FT	CONFLICT	158	158	T -> K (IN REF. 4).
FT	CONFLICT	184	184	R -> G (IN REF. 4).
FT	CONFLICT	215	215	S -> Y (IN REF. 4).
SO	SEQUENCE	316 AA;	35338 MW;	19140N/FD650CFB5 CRC64;

```

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RN SEQUENCE FROM N.A.
RA Issel-Tarver L., Rine J.:
RL Submitted (May-1996) to the EMBL/GenBank/DBJ databases.
RP SEQUENCE FROM N.A.
RA Bauer C., Williams D.:
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: PUTATIVE ODORANT RECEPTOR.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U56421; AAB01215.1; -
DR EMBL: AC004653; AAC64576.1; -
DR GCRDB: GCR_1924; -
DR InterPro: IPR000276; GPCR_Rhodopsn.
DR Pfam: PF00001; 7tm.1; 1.
DR PRINTS: PR00237; GPCRHOOPS.
DR PROSITE: PS00237; G-PROTEIN_RECEPT_F1_1; 1.
KW G-protein coupled receptor; Transmembrane; glycoprotein;
KW Multigene family; Olfaction.
FT DOMAIN 1 24 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 25 48 1 (POTENTIAL).
FT DOMAIN 49 57 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 58 79 2 (POTENTIAL).
FT DOMAIN 80 100 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 101 120 3 (POTENTIAL).
FT DOMAIN 121 139 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 140 160 4 (POTENTIAL).
FT DOMAIN 161 200 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 201 222 5 (POTENTIAL).
FT DOMAIN 223 236 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 237 261 6 (POTENTIAL).
FT DOMAIN 262 272 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 273 292 7 (POTENTIAL).
FT DOMAIN 293 317 CYTOPLASMIC (POTENTIAL).
FT DISULFID 97 189 BY SIMILARITY.
FT CARBOHYD 5 5 N-LINKED (GLCNAC...) (POTENTIAL).
FT CONFLICT 211 211 F -> L (IN REF. 1).
SQ SEQUENCE 317 AA; 35350 MW; A72ADE1AEB680CE CRC64;

Query Match 48.5%; Score 779; DB 1; Length 317;
Best Local Similarity 49.8%; Pred. No. 1.9e-51; Indels 0; Gaps 0;
Matches 150; Conservative 58; Mismatches 93;

QY 5 NDSIOEFILGFSRDMWLEFPLLVFLSYTYIFGNLTIIIVSRDTRLHTPMYFELT 64
DB 5 NQWVSEFILLGLSSDMDFRVSLFVFLVMVTVGNCILVILIRDSLHTPMYFELT 64
QY 65 NLSLLDLCYTCYVPMQVLVLCSTIRKVIYSGCAVQLFTFLAIGATEYLLLAWSFDMWY 124
DB 65 NLSLVDSYATSVVPMQLAFLAFLAHLKAIPFOCAAOGLFELFVLLAVMAYDRYV 124
QY 125 AICRPLHYVIMHORLCLOLAASWVGFNSWLSLTLOTPLCDPYVDHFLCEPAL 184
DB 125 AVCDALYTSALMHGGLCARLATISWSGFTSSPVQTAITPQLPCKRKFTDHTSCLELAV 184
QY 185 LKLSCVETTAENAEELFELFLPLTLILISYAIFVRAVLRIQSAEGKQAFKTCGSH 244
DB 185 VRLACVDTSSNEVTIMSSIVLMTPECVLVLSTIQIISITLIKQSGEKKKAFHRCASH 244

```

```

QY 245 LTVVSLFYSTAVSYLQPPSPSSKDGKMYSLFYGIAPMLNPLLYLRNKEVGEKRL 304
DB 245 LTVVSLCYGAIFTYIQPHSPSVYLQKLFVSFAILLPLMLNPATISLRNKEVGMQKL 304
QY 305 V 305
DB 305 L 305

RESULT 14
ID 05VL_HUMAN STANDARD; PRT; 321 AA.
AC 09UGF6;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE Olfactory receptor 5VL (Hs6M1-21).
GN OR5V1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RN SEQUENCE FROM N.A.
RA Tracey A.:
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: PUTATIVE ODORANT RECEPTOR.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AL096770; CAB65797.1; -
DR InterPro: IPR000276; GPCR_Rhodopsn.
DR Pfam: PF00001; 7tm.1; 1.
DR PRINTS: PR00237; GPCRHOOPS.
DR PROSITE: PS00237; G-PROTEIN_RECEPT_F1_1; 1.
DR PROSITE: PS50262; G-PROTEIN_RECEPT_F1_2; 1.
KW G-protein coupled receptor; Transmembrane; glycoprotein;
KW Multigene family; Olfaction.
FT DOMAIN 1 25 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 26 49 1 (POTENTIAL).
FT DOMAIN 50 57 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 58 79 2 (POTENTIAL).
FT DOMAIN 80 100 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 101 120 3 (POTENTIAL).
FT DOMAIN 121 139 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 140 156 4 (POTENTIAL).
FT DOMAIN 159 195 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 196 219 5 (POTENTIAL).
FT DOMAIN 220 236 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 237 259 6 (POTENTIAL).
FT DOMAIN 260 272 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 273 292 7 (POTENTIAL).
FT DOMAIN 293 321 CYTOPLASMIC (POTENTIAL).
FT DISULFID 97 189 BY SIMILARITY.
FT CARBOHYD 5 5 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 321 AA; 36056 MW; AAC426DFCE58E375 CRC64;

Query Match 47.7%; Score 767; DB 1; Length 321;
Best Local Similarity 49.3%; Pred. No. 1.5e-50; Indels 0; Gaps 0;
Matches 148; Conservative 53; Mismatches 99;

QY 5 NDSIOEFILGFSRDMWLEFPLLVFLSYTYIFGNLTIIIVSRDTRLHTPMYFELT 64
DB 5 NQWVSEFILLGLSSDMDFRVSLFVFLVMVTVGNCILVILIRDSLHTPMYFELT 64

```

```
QY 61 FFLTNLSLDLCYTTCTVPQMLVNLCSIRKVISYRGCAOLFPLALGATEYLLAAWMSF 120
    |||
DB 61 FFLTNLSLDLCYTTCTVPQMLINICSTRKVIISYGCVCVOLFPLISGSECEFLGWSL 120
QY 121 DMFPAICRPLHYVIMHQRCLQLAASWYTGFSNSWMLSTLTQLPLCDPYVIDHFLCE 180
    |||
DB 121 DRPLAICRPLHYVIMHQRCLQLAACWISGFSNSVLOSTWTLQMLPCGHKEVDHFCE 180
QY 181 VPALLKLSCEVTANAEALFLVSEFLPLILPLTLLISYAFIVRAVLRIOSEGRQKAFGT 240
    |||
DB 181 VPALLKLSCEVDITANAEALFLVSEFLPLILPLTLLISYAFIVRAVLRIOSEGRQKAFGT 240
QY 241 CGSHLIYVSLFYSTANSVYIQPPSPSSKDGKMWLSFYGIAPMLNPLIYLRNKEVKEG 300
    |||
DB 241 CGSHLIYVSLFYGTALYIMYIQPPSPSSKDGKMWLSFYGIITPMLNPLIYLRNEEVKGA 300
QY 301 FKRLVARVFLIKK 313
    |||
DB 301 FKRLMKRIILIKK 313
```

```
RESULT 2
ID 043883 PRELIMINARY; PRT; 216 AA.
AC 043883;
DT 01-JUN-1998 (TREMBLrel. 06, Created)
DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE OLFACTORY RECEPTOR (FRAGMENT).
GN OR5-40.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98160182; PubMed=9500546;
RA Rouquier S., Tavioux S., Trask B.J., Brand-Arpon V., van den Engh G.,
RA Demallie J., Giorli D.;
RT "Distribution of olfactory receptor genes in the human genome.";
RL Nat. Genet. 18:243-250(1998).
DR EMBL: U86270; AAC39628.1;
DR InterPro: IPR002106; AA_CRNA_ligase_II.
DR InterPro: IPR000276; GPCR_Rhodopsn.
DR Pfam: PF00001; 7tm_1; 1.
DR PROSITE: PS00339; AA_TRNA_LIGASE_II_2; UNKNOWN_1.
DR PROSITE: PS00237; G_PROTEIN_RECP_F1.1; UNKNOWN_1.
DR PROSITE: PS50262; G_PROTEIN_RECP_F1_2; 1.
KW Receptor.
FT NON_TER 1 1
FT NON_TER 216 216
SQ SEQUENCE 216 AA; 23963 MW; FF89E278BDIDA79F CRC64;
```

Query Match 68.3%; Score 1098; DB 4; Length 216;
Best Local Similarity 99.5%; Pred. No. 1e-93;
Matches 215; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
QY 68 LIDLCTTCVPMQVNLVCSIRKVISYRGCAOLFPLALGATEYLLAAWMSFDMWPAIC 127
    |||
DB 1 LIDLCTTCVPMQVNLVCSIRKVISYRGCAOLFPLALGATEYLLAAWMSFDMWPAIC 60
QY 128 RPLHYVIMHQRCLQLAASWYTGFSNSWMLSTLTQLPLCDPYVIDHFLCEVPALIKL 187
    |||
DB 61 RPLHYVIMHQRCLQLAASWYTGFSNSWMLSTLTQLPLCDPYVIDHFLCEVPALIKL 120
QY 188 SCVETTANAEALFLVSEFLPLILPLTLLISYAFIVRAVLRIOSEGRQKAFGTGSHLIY 247
    |||
DB 121 SCVETTANAEALFLVSEFLPLILPLTLLISYAFIVRAVLRIOSEGRQKAFGTGSHLIY 180
QY 248 VSLFYSTANSVYIQPPSPSSKDGKMWLSFYGIAP 283
    |||
```

DB 181 VSLFYSTANSVYIQPPSPSSKDGKMWLSFYGIAP 216

```
RESULT 3
ID 09GZK6 PRELIMINARY; PRT; 312 AA.
AC 09GZK6;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE OLFACTORY RECEPTOR
GN 6M1-4P*05 OR 6M1-4P*02.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Ziegler A., Ehlers A., Forbes S.A., Townsdales J., Uchanska-Ziegler B.,
RA Volz A., Younger R., Beck S.;
RT "Polymorphic olfactory receptor genes and HLA loci constitute extended
RT haplotypes.";
RL (In) Kasahara M. (eds.);
RL MAJOR HISTOCOMPATIBILITY COMPLEX-EVOLUTION, STRUCTURE, AND FUNCTION,
RL pp.110-130, Springer-Verlag, Tokyo, Japan (2000).
DR EMBL: AJ302570; CAC20490.1; -
DR EMBL: AJ302565; CAC20485.1; -
DR EMBL: AJ302566; CAC20486.1; -
DR EMBL: AJ302569; CAC20489.1; -
DR InterPro: IPR000276; GPCR_Rhodopsn.
DR Pfam: PF00001; 7tm_1; 1.
DR PRINTS: PR00237; GPCR_Rhodopsn.
DR PROSITE: PS00237; G_PROTEIN_RECP_F1.1; UNKNOWN_1.
DR PROSITE: PS50262; G_PROTEIN_RECP_F1_2; 1.
KW Receptor.
SQ SEQUENCE 312 AA; 35454 MW; 0C08C09B2DBA4835 CRC64;
```

Query Match 57.1%; Score 917; DB 4; Length 312;
Best Local Similarity 58.8%; Pred. No. 7.7e-77;
Matches 177; Conservative 43; Mismatches 81; Indels 0; Gaps 0;

```
QY 5 NDSITIDEFILLGSDRPMLEFPLLVFLISYTYTGNGNLITLLVSRIDTYLHPMWFELT 64
    |||
DB 6 NASFEDEFILLGFSNMWHELVFLVFLITLIGNELIILSYSDSHLHPMWFELS 65
QY 65 NLSLDLCYTTCTVPQMLVNLCSIRKVISYRGCAOLFPLALGATEYLLAAWMSFDMW 124
    |||
DB 66 NLSFLDLCYTTSSIPQMLVNLGPEKTIYAGCTVOLYFVLAIGTACVLLVWMSIDRIA 125
QY 125 AICRPLHYVIMHQRCLQLAASWYTGFSNSWMLSTLTQLPLCDPYVIDHFLCEVPAL 184
    |||
DB 126 AICRPLHYVIMHQRCLQLAASWYTGFSNSWMLSTLTQLPLCDPYVIDHFLCEVPAL 185
QY 185 LKLSCEVTANAEALFLVSEFLPLILPLTLLISYAFIVRAVLRIOSEGRQKAFGTGSH 244
    |||
DB 186 LKLSCEVTANAEALFLVSEFLPLILPLTLLISYAFIVRAVLRIOSEGRQKAFGTGSH 245
QY 245 LIYVSLFYSTANSVYIQPPSPSSKDGKMWLSFYGIAPMLNPLIYLRNKEVKEGFKRL 304
    |||
DB 246 LIYVSLFYSTANSVYIQPPSPSSKDGKMTALTYVTPLNPLIYLRNKEVKEGFKRL 305
QY 305 V 305
DB 306 M 306
```

```
RESULT 4
ID 09GZK1 PRELIMINARY; PRT; 312 AA.
AC 09GZK1;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
```

Query Match	Similarity	Score	ID	Length
Best Local	58.5%	Pred. No. 1.2e-76;		
Matches 176;	Conservative	44;	Mismatches 81;	Indels 0; Gaps 0;
OY	5	NDSIQIEFLLGLGSDRPMLEFPLLVLFVLSYVVTTFGNLTILVSRLDPTKHTPMYFELT	64	
DB	6	NASFDFDPLLGLGFSNMPHLEVVLFVLLITLIGNFIITLSYLDSSHLPMTFFLS	65	
OY	65	NISLIDLCTTCTVDPQMLVNLCSIRKVISYRCVAAQLFETALGATEYLLAVMSFDMFV	124	
DB	66	NLSFIDLCTYSSIPQLLVNLGMPEKRTISYACCTVQLVFALGTAECVLLVMSYDRYA	125	
OY	125	AICRPLHYSVIMHQRCLQGLAAASVTSNGSNVSLSTLTLOPLCDPYIDHFLCEVPL	184	
DB	126	ANCRPLHTVTLVLMHPFCRLTAAASVSGFTTSALHSSFTFWMPLCRHRVYDHFCEVPL	185	
OY	185	LKLSCEVTANAEAEFLVSEFLPLPLFLITLSYAFIVAVLRIOASAEGRKAFGCGSH	244	
DB	186	LRLSCVDQADQANLTLTMWSSIFVLLPLILITLSTGAIARAAYVSMOSTGLQGVLRTCGH	245	
OY	245	LIVSLFSTAVSYVLQPPSPSSKQDGKWSLVFGIILAPMLPLIYTLNNKEVKEGFKRL	304	
DB	246	LMVSLFLFIPVCMVLOPPSENSODQKFIALFVTVVPLNPILYTFERNKDVGVAKRL	305	
OY	305	V 305		
DB	306	M 306		
RESULT	5			
O9WV14				
ID	O9WV14	PRELIMINARY:	PRT:	310 AA.
AC	O9WV14:			
DT	01-NOV-1999	(TREMBLrel. 12, Created)		
DT	01-NOV-1999	(TREMBLrel. 12, Last sequence update)		
DT	01-JUN-2001	(TREMBLrel. 17, Last annotation update)		
DE	573K1.2 (MM17M1-3 (NOVEL 7 TRANSMEMBRANE RECEPTOR (RHODOPSIN FAMILY) (OPFACORY RECEPTOR LIKE) PROTEIN)).			
DE	(GABRI OR 573K1.2.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
NCBI	NCBI_TaxID=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Younger R.;			

Query Match	56.5%;	Score 908;	DB 11;	Length 310;
Best Local Similarity	57.9%;	Pred. No. 5.2e-76;		
Matches 175;	Conservative 50;	Mismatches 77;	Indels 0;	Gaps 0;

RESULT 6
Q921K0
ID Q921K0 PRELIMINARY; PRT; 302 AA

DT	01-MAY-1999 (TREMblrel_10, Created)
DT	01-MAY-1999 (TREMblrel_10, Last sequence update)
DT	01-DEC-2001 (TREMblrel_19, Last annotation update)
DE	OLFACTORY RECEPTOR 89 (FRAGMENT).
GN	OLFR89.
OS	Mus musculus (Mouse).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX	NCBI_TaxID=10090;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	TISSUE=EPITHELIUM;
RA	Ayer l'elievre C., Amadou C., Gallinaro H., Ayoustin P., Ribouchon M.,
RA	Bouissou C., Tazi Ahnani R., Pontarotti P.;
RT	"Olfactory receptor gene cluster in man and mouse major
RT	histocompatibility complex (MHC): New insights into the evolution of
RT	vertebrate olfactory receptor gene family."
RL	Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
DR	EMBL, AJ132195; CNA10603.1; -
DR	InterPro; IPR000276; GPCR_Rhodopsn.
DR	Pfam; PF00001; 7tm_1; 1.
DR	PRINTS; PR00237; GPCRHHODOPSN.
DR	PROSITE; PS00237; G_PROTEIN_RECPT_F1_1; UNKNOWN_1.
DE	PROSITE; PS50262; G_PROTEIN_RECPT_F1_2; 1.
KM	Receptor.
FT	NON_TER 1
SO	SEQUENCE 302 AA; 33971 MW; 19A4B7C3F1ADB0A CRC64;

Query Match	55.5%;	Score 892;	DB 11;	length 302;
Best Local Similarity	56.9%;	Pred. No. 1.5e-74;		
Matches 174;	Conservative 48;	Mismatches 76;	Indels 8;	Gaps 2

[illegible]

RESULT	7
09WV11	
ID	09WV11
AC	09WV11;
DT	01-NOV-1999 (TrEMBLrel. 12, Created)
DT	01-NOV-1999 (TrEMBLrel. 12, last sequence update)
DT	01-JUN-2001 (TrEMBLrel. 17, last annotation update)
DE	573k1.8 (MM17M1-2 (NOVEL 7 TRANSMEMBRANE RECEPTOR (RHODOPSIN FAMILY
DE	(OLEFACTORY RECEPTOR LIKE) PROTEIN)).
GN	GABBR1 OR 573K1.8.
OS	Mus musculus (Mouse).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
OX	NCBI_TaxID=10090;
LN	[1]
RP	SEQUENCE FROM N.A.
RA	Younger R.;
RL	Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
RL	EMBL: AI078630; CAB45000.1; -.
DR	MCD; MGI:1860139; Gabbrl.
DR	InterPro: IPR00276; GPCR_Rhodpsn.
DR	Pfam: PF00001; 7tm_1.1.
DR	PRINTS: PR00237; GPCRHHODPSN.
DR	PROSITE: PS00237; G_PROTEIN_RECP_FL1; UNKNOWN1.
DR	PROSITE: PS50262; G_PROTEIN_RECP_FL2; 1.
DR	Proteor: Transmembrane
QO	SEQUENCE 312 AA; 34786 MW; DAA7507C77710AD8 CRC64;

	Query Match	54.8%	Score 861;	DB 11,	Length 312;
	Best Local Similarity	55.6%,	Pred. No.1.6e-73;		
	Matches 169; Conservative	49;	Mismatches 86;	Indels 0;	Gaps 0.
QY	4 VNDSTIQEILLGSEDRPMLPEPLLVFLISTVTIVFGMLTIILYSRLDKLHTPYEFL	63			
	: : : : : : : : : : : :				
Dd	2 VNQSPVGFELLGFESEHPQLERVLVVVLCSTLLTGTLITLLSTDPRLSHPMYEFL	61			
QY	64 TNLSLDLCYTCTYPOMLVNCSIRKVISYGCAQAQLFIALGATEVELLAWSFDMF	123			
	: : : : : : : : : : : : :				
Dd	62 SNLSPLDLCEFTTVCPOMLFNLMGPRTKTSIFGSCSQVLTFEMLLTECILLTVMAFDREY	121			
QY	124 VAICRPMTSVYMHORCLQLAASAWVGFSNWSMLSTITLDLPICDPYVIDHPLCFEVNA	183			
	: : : : : : : : : : :				

D b	122	VAVQCP ^{LH} IYATIIHPRLCRQLAGVAMMIGLVGSIVQI ^P PTLLPFCSHRQIDDFLCEVPS	181
O y	184	LKIKSCVEETTANAEEDFLVSEELFHLIP ^T LTLLISYAFIYRAVLRTOSAGEKORAKFGTCGS	243
D b	182	LIRLSGGDTTFNEIOLSAVGIVFLFWPLSLTIYSYGVIARVA ^L TLNTSSKKRRNFYGCSS	241
O y	244	HLIYVSLEFYTAVASVYLQPPSPSSKDCKMWLSLFYGIIA ^R PLNPLIYTLLNKEEKGEGFR	303
D b	242	HLIYVTEFSYSSIVAYVLQPKNPYAQERSKFEGLFYAVGTP ^L LNPLVYTLRNKEVKRAFMR	301
O y	304	LVAR	307
D b	302	ILGK	305

RESULT	8	
09R02Z		
ID	09R02Z	PRELIMINARY;
AC	09R02Z;	PRF: 312 AA.
DT	01-MAY-2000 (TREMBLrel. 13, Created)	
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)	
DT	01-JUN-2001 (TREMBLrel. 17, Last annotation update)	
DE	573K1.3 (MM17M1-4 (NOVEL 7 TRANSMEMBRANE RECEPTOR (RHODOPSIN FAMILY) (OLFACTORY RECEPTOR LIKE) PROTEIN)).	
GN	GABBR1 OR 573K1.3.	
OS	Mus musculus (Mouse).	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC	Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.	
OX	NCBI_TaxID=10090;	
RN	[1]	
RA	SEQUENCE FROM N.A.	
RP	Younger R.;	
RL	Submitted (MAY-1999) to the EMBL/Genbank/DBJ databases.	
DR	EMBL: AL078630; CAB4495.1; -	
DR	MED; MG1:1860139; Gabbrl.	
DR	InterPro: IPR000276; GPCR_Rhodpsn.	
DR	Pfam: PF00001; 7tm_1.1.	
DR	PRINTS: PS00237; GPCRHHODPSN.	
DR	PROSITE: PS00237; G_PROTEIN_RECEP_FL1; UNKNOWN_1.	
DR	PROSITE: PS0262; G_PROTEIN_RECEP_FL2; 1.	
DR	Receptor; Transmembrane.	
QO	SEQUENCE 312 AA; 34999 MW; 83374B1A4C6F912D CRC64;	

	Query Match	Similarity	54.3%	Score 872;	DB 11;	length 312;
	Best Local	Similarity	55.6%	Pred. No. 1.le-72;		
	Matches 169;	Conservative	51;	Mismatches 84;	Indels 0;	Gaps 0;
QY	4	VNDSDIOEFILGFSDRPMLEFPPLVFLISYTVTFICGNLTIIIVSRDLTKLHTPMYFEL	63			
	1	: : : : : : : : : : :				
	2	VNOSPVGVEFLILGFSSEHQLEKLEFVILVCSTYLLTGNTLLTLLSTLIDPRLHSPMYFEL	61			
QY	64	TNLSLIDLCYTTCTYPPQMLVNLCSIRKVIISYGCVAQOLFPLALGATEYELLLAVNSPDMF	123			
	1	: : : : : : : : : :				
Db	62	SNLSFLDLCFTYTCVPOMLFNLGPAKTIISFGCSVOLEFISLDTETCILLTVMSFDXY	121			
	1	: : : : : : : : : :				
QY	124	VAICRPLHSYVIMHORLCLOLAASMTVGFSSNWSMLSTLLDLPICDDPYIDHIFCEYVA	183			
	1	: : : : : : : : : :				
Db	122	VAVOCPLHAYIYIHPRLCMKLAAVAMMGGLIOSIYQTPTLKLFPCCPHRQIDDFICEYVS	181			
	1	: : : : : : : : : :				
QY	184	LKLSCEVTETANAEELFVSELPHLIPTLILISYAFIVRAVLRIQSAEGROKAFGTCS	243			
	1	: : : : : : : : : :				
Db	182	LIRLSGDTTFNEIOLAVSSVILVYVPLSLILVSGALARAAMRLNSTEAMKMLRTGSS	241			
	1	: : : : : : : : : :				
QY	244	HLIVSLFYSYAVSVYIOLPPSPSSKADQCKNWSLEFGIILAPMLNPILYTLRNKEVGEGR	303			
	1	: : : : : : : : : :				
Db	242	HLIVTLEFYSYAVIAYLOPKNBYAOERBGKFEGLFYAVGAPPTLNPLIYTLRNKEVRAFEWR	301			
	1	: : : : : : : : : :				
QY	304	LYAR 307				
	1	:				
Db	302	LGK 305				


```
Db 1 MPTNESHPEEFILGADPRLWLEPLFTSLIMYPIAVNGNTIILMSRLDSRLSPMY 60
QY 61 FFLTNLSLDLCYTCVPMOMLVNLCIRKIVISYRGCAQLEFIALGATEYLLAAVMSF 120
Db 61 FFLTNLSLDLCYTCVPMOMLVNLCIRKIVISYRGCAQLEFIALGATEYLLAAVMSF 120
QY 121 DMFVAICRPLHYSVIMHORLCLOL--AAASW-----VYFNSVWLSLTTL 164
Db 121 DMFVAICRPLHYSVIMHORLCLOL--AAASW-----VYFNSVWLSLTTL 164
QY 165 QLELCPDYVDHFLCEYFALKLSCVETANEAEFLYSELPHILPTLLISTAFIVRA 224
Db 165 QLELCPDYVDHFLCEYFALKLSCVETANEAEFLYSELPHILPTLLISTAFIVRA 224
QY 225 VLRIOSAEGROKAFGTCGSHLIVSLFYSTAVSYIOLPPSSSDQCKWVSLFYGIATPM 284
Db 225 VLRIOSAEGROKAFGTCGSHLIVSLFYSTAVSYIOLPPSSSDQCKWVSLFYGIATPM 284
QY 285 LNPILYTLRNKEVEKEGFKRLVARY 308
Db 285 LNPILYTLRNKEVEKEGFKRLVARY 308
RESULT 12
Q92022 PRELIMINARY; PRT; 318 AA.
AC 092022;
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DE 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE B7 OLFACORY RECEPTOR.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;
RX MEDLINE=21310002; PubMed=11416212;
RA Lane R.P., Cutforth T., Young J., Athanasios M., Friedman C.,
RA Rowen L., Evans G., Axel R., Hood L., Trask B.J.;
RT "Genomic analysis of orthologous mouse and human olfactory receptor
RT locl.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:7390-7395(2001).
DR EMBL; AF321234; AAK84169.1; -.
KW Receptor.
SQ SEQUENCE 318 AA; 35150 MW; E9A182AEA21BSCAO CRC64;
```

```
QY 304 LVARVFLIK 312
Db 304 LVGRSLTVK 312
RESULT 13
Q9EPG5 PRELIMINARY; PRT; 314 AA.
ID Q9EPG5;
AC Q9EPG5;
DT 01-MAR-2001 (TReMBLrel. 16, Created)
DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)
DE B2 OLFACORY RECEPTOR.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129/SVJ;
RX MEDLINE=21310002; PubMed=11416212;
RA Lane R.P., Cutforth T., Young J., Athanasios M., Friedman C.,
RA Rowen L., Evans G., Axel R., Hood L., Trask B.J.;
RT "Genomic analysis of orthologous mouse and human olfactory receptor
RT locl.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:7390-7395(2001).
DR EMBL; AF321233; AAG45183.1; -.
DR InterPro: IPR000276; GPCR_Rhodopsn.
DR InterPro: IPR000478; TNF_family.
DR Pfam: PF00001; 7tm_1; 1.
DR PRINTS: PR00237; GPCR_RHODOPSIN.
DR PROSITE: PS50262; G_PROTEIN_RECPT_FL_2; 1.
DR PROSITE: PS00251; TNF_1; UNKNOWN_1.
KW Receptor.
SQ SEQUENCE 314 AA; 35154 MW; 6325415D032FCD1B CRC64;
```

```
Query Match 49.9%; Score 802; DB 11; Length 318;
Best Local Similarity 51.1%; Pred. No. 3.3e-66;
Matches 158; Conservative 58; Mismatches 91; Indels 2; Gaps 2;
```

```
Query Match 49.3%; Score 793; DB 11; Length 314;
Best Local Similarity 51.8%; Pred. No. 2.2e-65;
Matches 158; Conservative 54; Mismatches 91; Indels 2; Gaps 1;
```


GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 25, 2002, 18:24:47 ; Search time 3373.42 Seconds
(without alignments)
6047.782 Million cell updates/sec

Title: US-09-800-321A-3
Perfect score: 943
Sequence: 1 tatgaattggtaataagaca.....tcttcttaatacaagaataaa 943

Scoring table:
IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 21979536 seqs, 10817449327 residues

Total number of hits satisfying chosen parameters: 43959072

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 08
Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents_NA_Main:*
1: /cgn2_6/ptodata/2/pna/US006.COMB.seq:*
2: /cgn2_6/ptodata/2/pna/US006.COMB.seq:*
3: /cgn2_6/ptodata/2/pna/US007.COMB.seq:*
4: /cgn2_6/ptodata/2/pna/US008.COMB.seq:*
5: /cgn2_6/ptodata/2/pna/US008.COMB.seq:*
6: /cgn2_6/ptodata/2/pna/US008.COMB.seq:*
7: /cgn2_6/ptodata/2/pna/US008.COMB.seq:*
8: /cgn2_6/ptodata/2/pna/US008.COMB.seq:*
9: /cgn2_6/ptodata/2/pna/US008.COMB.seq:*
10: /cgn2_6/ptodata/2/pna/US008.COMB.seq:*
11: /cgn2_6/ptodata/2/pna/US008.COMB.seq:*
12: /cgn2_6/ptodata/2/pna/US008.COMB.seq:*
13: /cgn2_6/ptodata/2/pna/US008.COMB.seq:*
14: /cgn2_6/ptodata/2/pna/US009.COMB.seq:*
15: /cgn2_6/ptodata/2/pna/US009.COMB.seq:*
16: /cgn2_6/ptodata/2/pna/US009.COMB.seq:*
17: /cgn2_6/ptodata/2/pna/US009.COMB.seq:*
18: /cgn2_6/ptodata/2/pna/US009.COMB.seq:*
19: /cgn2_6/ptodata/2/pna/US009.COMB.seq:*
20: /cgn2_6/ptodata/2/pna/US009.COMB.seq:*
21: /cgn2_6/ptodata/2/pna/US009.COMB.seq:*
22: /cgn2_6/ptodata/2/pna/US009.COMB.seq:*
23: /cgn2_6/ptodata/2/pna/US009.COMB.seq:*
24: /cgn2_6/ptodata/2/pna/US009.COMB.seq:*
25: /cgn2_6/ptodata/2/pna/US009.COMB.seq:*
26: /cgn2_6/ptodata/2/pna/US009.COMB.seq:*
27: /cgn2_6/ptodata/2/pna/US009.COMB.seq:*
28: /cgn2_6/ptodata/2/pna/US009.COMB.seq:*
29: /cgn2_6/ptodata/2/pna/US009.COMB.seq:*
30: /cgn2_6/ptodata/2/pna/US009.COMB.seq:*
31: /cgn2_6/ptodata/2/pna/US009.COMB.seq:*
32: /cgn2_6/ptodata/2/pna/US009.COMB.seq:*
33: /cgn2_6/ptodata/2/pna/US009.COMB.seq:*
34: /cgn2_6/ptodata/2/pna/US009.COMB.seq:*
35: /cgn2_6/ptodata/2/pna/US009.COMB.seq:*
36: /cgn2_6/ptodata/2/pna/US009.COMB.seq:*
37: /cgn2_6/ptodata/2/pna/US009.COMB.seq:*
38: /cgn2_6/ptodata/2/pna/US009.COMB.seq:*
39: /cgn2_6/ptodata/2/pna/US009.COMB.seq:*
40: /cgn2_6/ptodata/2/pna/US009.COMB.seq:*
41: /cgn2_6/ptodata/2/pna/US009.COMB.seq:*
42: /cgn2_6/ptodata/2/pna/US009.COMB.seq:*
43: /cgn2_6/ptodata/2/pna/US009.COMB.seq:*

Result				SUMMARIES				Description			
No.	Score	Query	Length	ID	DB	Score	Length	ID	DB	Score	Length
1	943	100.0	943	31	US-09-800-321A-3	943	31	US-09-800-321A-3	31	Sequence 3, Appl1	
2	940.4	99.7	993	31	US-09-800-321A-1	940.4	99.7	US-09-800-321A-1	31	Sequence 1, Appl1	
3	940.4	99.7	113710	64	US-60-258-250-40	940.4	99.7	US-60-258-250-40	64	Sequence 40, Appl1	
4	938.8	99.6	942	29	US-09-755-017-1	938.8	99.6	US-09-755-017-1	29	Sequence 1, Appl1	
5	937.2	99.4	1488	29	US-09-755-017-3	937.2	99.4	US-09-755-017-3	29	Sequence 3, Appl1	
6	937.2	99.4	956	31	US-09-800-321A-5	937.2	99.4	US-09-800-321A-5	31	Sequence 5, Appl1	
7	716.8	76.0	20410	65	US-60-261-976-77	716.8	76.0	US-60-261-976-77	65	Sequence 77, Appl1	
8	665.2	70.5	1047	75	US-60-360-207-43515	665.2	70.5	US-60-360-207-43515	75	Sequence 43515, A	
9	602.2	63.9	1071	75	US-60-360-207-42178	602.2	63.9	US-60-360-207-42178	75	Sequence 604, Appl	
10	600.4	63.7	673	56	US-60-178-308-604	600.4	63.7	US-60-178-308-604	56	Sequence 86, Appl	
11	534	56.6	46924	63	US-60-245-225-86	534	56.6	US-60-245-225-86	63	Sequence 274, Appl	
12	534	56.6	46924	64	US-60-258-250-27	534	56.6	US-60-258-250-27	64	Sequence 27, Appl	
13	534	56.6	68686	63	US-60-248-505-274	534	56.6	US-60-248-505-274	63	Sequence 1606, Ap	
14	529.4	56.1	1510	63	US-60-162-248-93	529.4	56.1	US-60-162-248-93	63	Sequence 93, Appl	
15	505.6	53.6	642	55	US-60-162-356-86	505.6	53.6	US-60-162-356-86	55	Sequence 86, Appl	
16	505.6	53.6	642	55	US-60-162-356-86	505.6	53.6	US-60-162-356-86	55	Sequence 238, App	
17	505.6	53.6	642	55	US-60-162-356-86	505.6	53.6	US-60-162-356-86	55	Sequence 1245, Ap	
18	505.6	53.6	642	55	US-60-162-356-86	505.6	53.6	US-60-162-356-86	55	Sequence 697, App	
19	486.4	51.6	488	56	US-60-178-308-697	486.4	51.6	US-60-178-308-697	56	Sequence 556, App	
20	484.4	47.6	1324	63	US-60-245-225-556	484.4	47.6	US-60-245-225-556	63	Sequence 167, App	
21	448.4	47.6	1324	64	US-60-258-250-167	448.4	47.6	US-60-258-250-167	64	Sequence 111618, A	
22	448	47.5	2940	24	US-09-634-3066-111618	448	47.5	US-09-634-3066-111618	24	Sequence 484, App	
23	426	45.2	438	57	US-60-186-652-484	426	45.2	US-60-186-652-484	57	Sequence 3, Appl1	
24	406.8	43.1	1037	34	US-09-924-359-3	406.8	43.1	US-09-924-359-3	34	Sequence 26, Appl1	
25	406.8	43.1	1037	34	US-09-924-359-3	406.8	43.1	US-09-924-359-3	34	Sequence 26, Appl1	
26	403.6	42.8	939	1	PCT-US01-07771-26	403.6	42.8	PCT-US01-07771-26	1	Sequence 26, Appl1	
27	403.6	42.8	939	1	PCT-US01-20122-26	403.6	42.8	PCT-US01-20122-26	1	Sequence 26, Appl1	
28	403.6	42.8	939	31	US-09-804-291-26	403.6	42.8	US-09-804-291-26	31	Sequence 26, Appl1	
29	403.6	42.8	939	33	US-09-886-055-26	403.6	42.8	US-09-886-055-26	33	Sequence 26, Appl1	
30	403.6	42.8	939	57	US-60-188-914-26	403.6	42.8	US-60-188-914-26	57	Sequence 1, Appl1	
31	403.6	42.8	998	34	US-09-924-359-1	403.6	42.8	US-09-924-359-1	34	Sequence 1, Appl1	

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Yu, M.
09/1800321
Seq. IDs 34 w/
Interf

32	402	12.6	939	59	US-60-360-222-19	Sequence 19, Appl
33	383.2	40.6	948	75	US-60-366-207-43111	Sequence 43111, A
34	371.6	39.4	948	26	US-09-667-889-4	Sequence 4, Appl11
35	371.6	39.4	948	36	US-09-978-825-4	Sequence 4, Appl11
36	371.6	39.4	1005	37	US-10-034-842-1	Sequence 1, Appl11
37	371.6	39.4	1500	37	US-10-034-842-3	Sequence 3, Appl11
38	370	39.2	971	34	US-09-912-976-19	Sequence 19, Appl
39	370	39.2	1005	1	PCT-US01-07771-366	Sequence 366, App
40	370	39.2	1005	1	PCT-US01-20132-366	Sequence 366, App
41	370	39.2	1005	31	US-09-804-291-366	Sequence 366, App
42	370	39.2	1005	33	US-09-886-055-366	Sequence 366, App
43	370	39.2	1065	19	US-09-524-730-7	Sequence 7, Appl11
44	370	39.2	1065	37	US-10-028-521-7	Sequence 7, Appl11
45	367.2	38.9	706	65	US-60-261-976-387	Sequence 387, App

ALIGNMENTS

```

RESULT 1
US-09-800-321A-3
: Sequence 3, Application US/09800321A
: GENERAL INFORMATION:
: APPLICANT: Padigaru, Muralidhara
: APPLICANT: Burgess, Catherine E
: APPLICANT: Mishra, Vishnu
: APPLICANT: Li, Li
: APPLICANT: Baumgartner, Jason C
: APPLICANT: Majumdar, Kumud
: APPLICANT: Spytek, Kimberly A
: APPLICANT: Tchiernev, Velizar T
: TITLE OF INVENTION: Novel Proteins and Nucleic Acids Encoding Same
: FILE REFERENCE: 15966-703 US
: CURRENT APPLICATION NUMBER: US/09/800,321A
: CURRENT FILING DATE: 2001-03-05
: PRIOR APPLICATION NUMBER: 60/186,606
: PRIOR FILING DATE: 2000-03-03
: PRIOR APPLICATION NUMBER: 60/221,942
: PRIOR FILING DATE: 2000-07-31
: PRIOR APPLICATION NUMBER: 60/260,285
: PRIOR FILING DATE: 2001-01-08
: PRIOR APPLICATION NUMBER: 60/220,263
: PRIOR FILING DATE: 2000-07-24
: PRIOR APPLICATION NUMBER: 60/257,600
: PRIOR FILING DATE: 2000-12-21
: PRIOR APPLICATION NUMBER: 60/187,295
: PRIOR FILING DATE: 2000-03-06
: PRIOR APPLICATION NUMBER: 60/219,854
: PRIOR FILING DATE: 2000-07-21
: PRIOR APPLICATION NUMBER: 60/187,249
: PRIOR FILING DATE: 2000-03-06
: PRIOR APPLICATION NUMBER: 60/187,247
: PRIOR FILING DATE: 2000-03-06
: PRIOR APPLICATION NUMBER: 60/187,250
: PRIOR FILING DATE: 2000-03-06
: PRIOR APPLICATION NUMBER: 60/187,253
: PRIOR FILING DATE: 2000-03-06
: PRIOR APPLICATION NUMBER: 60/187,248
: PRIOR FILING DATE: 2000-03-06
: PRIOR APPLICATION NUMBER: 60/187,296
: PRIOR FILING DATE: 2000-03-06
: PRIOR APPLICATION NUMBER: 60/187,563
: PRIOR FILING DATE: 2000-03-07
: NUMBER OF SEQ ID NOS: 78
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 3
: LENGTH: 943
: TYPE: DNA
: ORGANISM: Homo sapiens
US-09-800-321A-3

```

Query Match 100.0%; Score 943; DB 31; Length 943;

Best Local Similarity 100.0%; Pred. No. 1.4e-273;
Matches 943; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

[illegible]

RESULT 2
US-09-800-321A-1
; Sequence 1, Application US/09800321A
; GENERAL INFORMATION:
; APPLICANT: Padigaru, Muralidhara

```

: APPLICANT: Burgess, Catherine E
: APPLICANT: Mishra, Vishnu
: APPLICANT: Li, Li
: APPLICANT: Baumgartner, Jason C
: APPLICANT: Majumder, Kumud
: APPLICANT: Spytek, Kimberly A
: APPLICANT: Tchenerov, Velizar T
: TITLE OF INVENTION: Novel Proteins and Nucleic Acids Encoding Same
: FILE REFERENCE: 15966-703 US
: CURRENT APPLICATION NUMBER: US/09/800,321A
: CURRENT FILING DATE: 2001-03-05
: PRIOR APPLICATION NUMBER: 60/186,606
: PRIOR FILING DATE: 2000-03-03
: PRIOR APPLICATION NUMBER: 60/221,942
: PRIOR FILING DATE: 2000-07-31
: PRIOR APPLICATION NUMBER: 60/260,285
: PRIOR FILING DATE: 2001-01-08
: PRIOR APPLICATION NUMBER: 60/220,263
: PRIOR FILING DATE: 2000-07-24
: PRIOR APPLICATION NUMBER: 60/257,600
: PRIOR FILING DATE: 2000-12-21
: PRIOR APPLICATION NUMBER: 60/187,295
: PRIOR FILING DATE: 2000-03-06
: PRIOR APPLICATION NUMBER: 60/219,854
: PRIOR FILING DATE: 2000-07-21
: PRIOR APPLICATION NUMBER: 60/187,249
: PRIOR FILING DATE: 2000-03-06
: PRIOR APPLICATION NUMBER: 60/187,247
: PRIOR FILING DATE: 2000-03-06
: PRIOR APPLICATION NUMBER: 60/187,250
: PRIOR FILING DATE: 2000-03-06
: PRIOR APPLICATION NUMBER: 60/187,253
: PRIOR FILING DATE: 2000-03-06
: PRIOR APPLICATION NUMBER: 60/187,248
: PRIOR FILING DATE: 2000-03-06
: PRIOR APPLICATION NUMBER: 60/187,296
: PRIOR FILING DATE: 2000-03-06
: PRIOR APPLICATION NUMBER: 60/187,563
: PRIOR FILING DATE: 2000-03-07
: NUMBER OF SEQ ID NOS: 78
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 1
: LENGTH: 993
: TYPE: DNA
: ORGANISM: Homo sapiens
: US-09-800-321A-1

Query Match          99.7%: Score 940.4; DB 31; Length 993;
Best Local Similarity 99.9%: Pred. No. 8.8e-273;
Matches 941; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 atgaattggtgaatgacagcatcatcagagatttattctgctgggtttctcagatcga 61
    |||||||
DB 15 atgatttgggtaaatgacagcatcatcagagatttattctgctgggtttctcagatcga 74
    |||||||

QY 62 ccttgctgaggttccacctcctgtgtgtcttcttgatttcttaoactgacacatctt 121
    |||||||
DB 75 ccttgctgaggttccacctcctgtgtgtcttcttgatttcttaoactgacacatctt 134
    |||||||

QY 122 ggaactcggacattattctagtgtcaagcctggaacacaaacttacaacccatgat 181
    |||||||
DB 135 ggaactcggacattattctagtgtcaagcctggaacacaaacttacaacccatgat 194
    |||||||

QY 182 ttcttcttaacatctatcactcctggtatcttgttacaacacatgtagtccaca 241
    |||||||
DB 195 ttcttcttaacatctatcactcctggtatcttgttacaacacatgtagtccaca 254
    |||||||

QY 242 atgctagtaaatattatgacagcatcagaagaatcaatcagttatcgttgctgtgtagccca 301
    |||||||
DB 255 atgctagtaaatattatgacagcatcagaagaatcaatcagttatcgttgctgtgtagccca 314
    |||||||

QY 302 ctttcaatattctggccttggggctactgaatactctccctggcgccatgtccctt 361
    |||||||

```

```

DB 315 ctttcaatattctggccttggggctactgaatactctccctggcgccatgtccctt 374
    |||||||
QY 362 gattgttgtagcatttgcgcctccctccattaccatgattatcatgacacagagactc 421
    |||||||
DB 375 gatagttttagcatttgcgcctccctccattaccatgattatcatgacacagagactc 434
    |||||||

QY 422 tgcctcagttgagcagctgcatcctggtgttactggttttagtaacctcagtggtgtct 481
    |||||||
DB 435 tgcctcagttgagcagctgcatcctggtgttactggttttagtaacctcagtggtgtct 494
    |||||||

QY 482 accctgacttccacagctgcatcctggtgttactggttttagtaaccttctctgtgaa 541
    |||||||
DB 495 accctgacttccacagctgcatcctggtgttactggttttagtaaccttctctgtgaa 554
    |||||||

QY 542 gtccctgacgtgcctcaaatcttctgtgtttagaacaagaagaatgagctggaacttcc 601
    |||||||
DB 555 gtccctgacgtgcctcaaatcttctgtgtttagaacaagaagaatgagctggaacttcc 614
    |||||||

QY 602 ctgtcagtgagctcttccatctaataccctgacactcatcctatatacatatgcttt 661
    |||||||
DB 615 ctgtcagtgagctcttccatctaataccctgacactcatcctatatacatatgcttt 674
    |||||||

QY 662 attgtccagcagatattgagatagctgtcgtgaaggtgcacaaaagcatttggagca 721
    |||||||
DB 675 attgtccagcagatattgagatagctgtcgtgaaggtgcacaaaagcatttggagca 734
    |||||||

QY 722 tctgttcccatctaatgtgtgtctcttctttaaagcagccgtctcgtgtgactg 781
    |||||||
DB 735 tctgttcccatctaatgtgtgtctcttctttaaagcagccgtctcgtgtgactg 794
    |||||||

QY 782 caaccacctgcgcagctcctcaaggaagcaagaatgttctctctctatgaaatc 841
    |||||||
DB 795 caaccacctgcgcagctcctcaaggaagcaagaatgttctctctctatgaaatc 854
    |||||||

QY 842 attgacacctgctggaatcccttatatacttaagaaacaggaagtaaggaagc 901
    |||||||
DB 855 attgacacctgctggaatcccttatatacttaagaaacaggaagtaaggaagc 914
    |||||||

QY 902 tttaaaggtgtgtgcaagagcttcttctaatacaagaataa 943
    |||||||
DB 915 tttaaaggtgtgtgcaagagcttcttctaatacaagaataa 956
    |||||||

RESULT 3
US-60-258-250-40/c
: Sequence 40, Application US/60258250
: GENERAL INFORMATION:
: APPLICANT: Beasley, Ellen
: TITLE OF INVENTION: ISOLATED HUMAN G-PROTEIN COUPLED
: TITLE OF INVENTION: RECEPTORS, NUCLEIC ACID MOLECULES ENCODING HUMAN GPCR
: FILE REFERENCE: CLO01035-PROV
: CURRENT APPLICATION NUMBER: US/60/258,250
: CURRENT FILING DATE: 2000-12-27
: NUMBER OF SEQ ID NOS: 210
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 40
: LENGTH: 113710
: TYPE: DNA
: ORGANISM: HUMAN
: US-60-258-250-40

Query Match          99.7%: Score 940.4; DB 64; Length 113710;
Best Local Similarity 99.9%: Pred. No. 8.6e-272;
Matches 941; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 atgaattggtgaatgacagcatcatcagagatttattctgctgggtttctcagatcga 61
    |||||||
DB 91948 ATGAATTGGGTAAATGACAGCATCATACAGAGATTATTCGTGCTGGTTCTCAGATCGA 91889
    |||||||

QY 62 ccttgctgaggttccacctcctgtgtgtcttcttgatttcttaoactgacacatctt 121
    |||||||

```

```
Db 91868 CTTGGCTGGAGTTTCCACTCTTGTGGTCTTGTGATTTCTTACACGTGACCAATCTTT 91829
Qy 122 ggcacatcgacatattatctagtcgacgcctcgagacacacatccacccatgcat 181
Db 91828 GGCATATCGACATATTATCTGATGTCAGCGCTCGACACCAACTTCATACCCCATGAT 91769
Qy 182 tttttcttacaatctacacccctcgatcttctgttacaacaatgatacagtcacaa 241
Db 91768 TTTTCTTACCAATCTATCTACCTCGATCTTGTGTACACCACTGATGATGCCACAA 91709
Qy 242 atgcagagaaattatgacagatcaagaaataatcagttatcgtggtgtgtgtgagccag 301
Db 91708 ATGCTAGTAATTTATGACGATCGAGAAATATCACTATATCGTGCTGTGAGCCAG 91649
Qy 302 ctcttcaattctcgcccttgaggagctactgaatactctccctcgagccgcatgctctt 361
Db 91648 CTTTTCATATTTCTGCGCTTGGGGCTACTGAATATCTTCTCCGCGCGCTCATGCTCTT 91589
Qy 362 gattggtttgtagctattgttcgagcctctcattactcaatgatacagacagagactc 421
Db 91588 GATAGGTTTGTAGCTATTTGTGCGGCTCTCCATTACTCAGTTATCATGACACAGACTC 91529
Qy 422 tgcctcagttgagagcgcacactcctggttactggttttagtaactaagtggtgtct 481
Db 91528 TGCCCTCCAGTTGGAGCTGCAATCCCTGGGTACTGGTTTACTGATCACTGATGCTGTCT 91469
Qy 482 accctgacatccagctcgacactcctgttaccctatgtagatcaattctcctgtgaa 541
Db 91468 ACCCGACTCCAGCTCGACGCTCCACTCGTGACCCCTATGATGATGATCACTTCTCTGTAA 91409
Qy 542 gtccctgacatgctcaagttactctgtgtgtgagacaagaacaaatgagctgaactatc 601
Db 91408 GTCCCTGCACGTCCCAATTTCTTGTGTGAGACAAACAGCAAAATGAGGCTGAACTATTC 91349
Qy 602 ctctcagtagagctcttccatctaataccctcgacacacacatctatactatgcttt 661
Db 91348 CTGTGACGTAGCTCTTTCATCTAATATACCCCTGACACTCATCTTATATCATATGCTTT 91289
Qy 662 atgtccagagagatitgagagatacagctctgaaagtcgacaaaaagacattcgagaca 721
Db 91288 ATGTGCCAGCAGATTTAGATGATACGCTGCTGTAAGCTGCACAAAGCATTTGGGACA 91229
Qy 722 tctgtgtcccatcaattgtgtgtctcttttatactagtagacgcgctctctgtacctg 781
Db 91228 TGTGTTTCCCATCTAATTTGTGTGCTCTTTTATATACAGCCGCTCTGTGATCCTG 91169
Qy 782 caaccacctcgagcagctcgaagacgaagaagaatggttctctctctatggaatc 841
Db 91168 CAACCACTTGGCCAGCTCCAGAGACCAAGAAAGATGTTTCTCTCTATGGAATC 91109
Qy 842 atgcaaccatgctgaatccctctatatatacacttaggaacaagaagtagtaagaagc 901
Db 91108 ATGACACCCAGTGTGAATCCCTTATATATACCTTAGAACAAGAGAGTAAGAAGC 91049
Qy 902 tttaaaagttgtgtgcaagaagcttcttaatacaagaataa 943
Db 91048 TTTAAAGTTGTGTGCAAGAGCTTCTTATCAAGAAATAA 91007
```

```
RESULT 4
US-09-755-017-1
; Sequence 1, Application US/09755017
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Milganowski, Nathaniel
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Friedlich, Glenn
; APPLICANT: Abulin, Alejandro
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: Novel Human Membrane Proteins and
; TITLE OF INVENTION: Polynucleotides Encoding the Same
```

```
FILE REFERENCE: LEX-0115-USA
; CURRENT APPLICATION NUMBER: US/09/755,017
; CURRENT FILING DATE: 2001-01-05
; PRIOR APPLICATION NUMBER: US 60/175,764
; PRIOR FILING DATE: 2000-01-12
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO: 1
; LENGTH: 942
; TYPE: DNA
; ORGANISM: Homo Sapien
US-09-755-017-1

Query Match 99.6%; Score 938.8; DB 29; Length 942;
Best Local Similarity 99.8%; Pred. No. 2.6e-272;
Matches 940; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 atgaatggtgtaaatgacagcatcatcagagagttatctcgtcgggttctcagatcga 61
Db 1 atgaatggtgtaaatgacagcatcatcagagagttatctcgtcgggttctcagatcga 60
Qy 62 ccttgagctgaggttccacactcctggtgtctcttgaatttcttaactgtgagacactctt 121
Db 61 ccttgagctgaggttccacactcctggtgtctcttgaatttcttaactgtgagacactctt 120
Qy 122 ggcacatcgacatattatctagtcgacgcctcgagacacacacatccacatgcat 181
Db 121 ggcacatcgacatattatctagtcgacgcctcgagacacacacatccacatgcat 180
Qy 182 tttttcttacaatctacacccctcgatcttctgttacaacaatgatacagtcacaa 241
Db 181 tttttcttacaatctacacccctcgatcttctgttacaacaatgatacagtcacaa 240
Qy 242 atgcagagaaattatgacagatcaagaaataatcagttatcgtgtgtgtgtgagccag 301
Db 241 atgcagagaaattatgacagatcaagaaataatcagttatcgtgtgtgtgtgagccag 300
Qy 302 ctcttcaattctcgcccttgaggagctactgaatactctcctcgagccgcatgctctt 361
Db 301 ctcttcaattctcgcccttgaggagctactgaatactctcctcgagccgcatgctctt 360
Qy 362 gattggtttgtagctattgttcgagcctctcattactcaatgatacagacagagactc 421
Db 361 gattggtttgtagctattgttcgagcctctcattactcaatgatacagacagagactc 420
Qy 422 tgcctcagttgagcagctcctggtgttactggttttagtaactaagtggtgtct 481
Db 421 tgcctcagttgagcagctcctggtgttactggttttagtaactaagtggtgtct 480
Qy 482 accctgacatccagctgacactctgttgaaccctatgtatagataccaattctctgtgaa 541
Db 481 accctgacatccagctgacactctgttgaaccctatgtatagataccaattctctgtgaa 540
Qy 542 gtccctgacatgctcaagttactctgtgtgtgagacaagaagaatgagcgtgaactatc 601
Db 541 gtccctgacatgctcaagttactctgtgtgtgagacaagaagaatgagcgtgaactatc 600
Qy 602 ctgtcagtagagctcttccatctaatccccctgagacactcattatatacagcttt 661
Db 601 ctgtcagtagagctcttccatctaatccccctgagacactcattatatacagcttt 660
Qy 662 atgtccagagagatitgagagatacagctctgttgaagtcgacaaaaagacattgagaca 721
Db 661 atgtccagagagatitgagagatacagctctgttgaagtcgacaaaaagacattgagaca 720
Qy 722 tctgtgtcccatcaattgtgtgtctcttttatactagtagacagcgtctctgtacctg 781
Db 721 tctgtgtcccatcaattgtgtgtctcttttatactagtagacagcgtctctgtacctg 780
Qy 782 caaccacactcgccagctcgaagacgaagaagaatggtttctctctctatggaatc 841
Db 781 caaccacactcgccagctcgaagacgaagaagaatggtttctctctctatggaatc 840
```

QY 842 attgcacccatgtcgaatcccttataatacaactaggacaaggagttaagaagc 901
 |||||||
 DB 841 atgcacccatgtcgaatcccttataatacaactaggacaaggagttaagaagc 900
 QY 902 tttaaaaggttggttcgaagagcttcttcaatcaagaataa 943
 |||||||
 DB 901 tttaaaaggttggttcgaagagcttcttcaatcaagaataa 942

RESULT 5
 US-09-755-017-3

Sequence 3, Application US/09755017
 GENERAL INFORMATION:
 APPLICANT: Walke, D. Wade
 APPLICANT: Wilganowski, Nathaniel
 APPLICANT: Turner, C. Alexander Jr.
 APPLICANT: Friedrich, Glenn
 APPLICANT: Abuin, Alejandro
 APPLICANT: Zambrowicz, Brian
 APPLICANT: Sands, Arthur T.
 TITLE OF INVENTION: Novel Human Membrane Proteins and
 TITLE OF INVENTION: Polynucleotides Encoding the Same
 FILE REFERENCE: LEX-0115-USA
 CURRENT APPLICATION NUMBER: US/09/755,017
 PRIOR FILING DATE: 2001-01-05
 PRIOR APPLICATION NUMBER: US 60/175,764
 PRIOR FILING DATE: 2000-01-12
 NUMBER OF SEQ ID NOS: 3
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 3
 LENGTH: 1488
 TYPE: DNA
 ORGANISM: Homo Sapien
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: (1)..(1488)
 OTHER INFORMATION: n = A,T,C or G
 US-09-755-017-3

Query Match 99.6%; Score 938.8; DB 29; Length 1488;
 Best Local Similarity 99.8%; Pred. No. 3.3e-272;
 Matches 940; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 2 atgaattgggttaaatgacagcatalacagaggttatctctgtggttctcagatcga 61
 |||||||
 DB 198 atgaattgggttaaatgacagcatalacagaggttatctctgtggttctcagatcga 257
 QY 62 ccttgctgaggtttccactccttctgtgtcttcttgaattcttcaactgtgacatctt 121
 |||||||
 DB 258 ccttgctgaggtttccactccttctgtgtcttcttgaattcttcaactgtgacatctt 317
 QY 122 ggcacatcacatattatagtgatcagccttgagacacacaaattatatacccatgat 181
 |||||||
 DB 318 ggcacatcacatattatagtgatcagccttgagacacacaaattatatacccatgat 377
 QY 182 tttttcttcaacatcatcactcctgtgattctgtttaacacacatgatacgtccacaa 241
 |||||||
 DB 378 tttttcttcaacatcatcactcctgtgattctgtttaacacacatgatacgtccacaa 437
 QY 242 atgctatgataattatgacagcatcaagaaagtaacagtatactgtgtgtagccag 301
 |||||||
 DB 438 atgctatgataattatgacagcatcaagaaagtaacagtatactgtgtgtagccag 497
 QY 302 ctttcaattttcgtccttgagggttactgaatcttctctcgtgcgcatgctctt 361
 |||||||
 DB 498 ctttcaattttcgtccttgagggttactgaatcttctctcgtgcgcatgctctt 557
 QY 362 gatgtgttgtagctattgttcgaccttcacattactcagttatcagtcagcagagactc 421
 |||||||
 DB 558 gatgtgttgtagctattgttcgaccttcacattactcagttatcagtcagcagagactc 617

QY 422 tgcctcagttggtcagctgcatactctgtgttaactggttttaactcagtggtgtct 481
 |||||||
 DB 618 tgcctcagttggtcagctgcatactctgtgttaactggttttaactcagtggtgtct 677
 QY 482 accctgactcagctgcagcactctgtgacctatgtgaatgaatcaactctctgtgaa 541
 |||||||
 DB 678 accctgactcagctgcagcactctgtgacctatgtgaatgaatcaactctctgtgaa 737
 QY 542 gtccctgacatgtctcaagtatctgtgttgagacaacaagcaaatgaggctgaacttc 601
 |||||||
 DB 738 gtccctgacatgtctcaagtatctgtgttgagacaacaagcaaatgaggctgaacttc 797
 QY 602 ctgtgtagtgagcttcttcaatcttaataccctgaactatccttataatattgtctt 661
 |||||||
 DB 798 ctgtgtagtgagcttcttcaatcttaataccctgaactatccttataatattgtctt 857
 QY 662 attgcacagcagtatgtgaagatacagctgcctggaagtgacaaaagcatttgagaca 721
 |||||||
 DB 858 attgcacagcagtatgtgaagatacagctgcctggaagtgacaaaagcatttgagaca 917
 QY 722 tgtgttcccatctaatgtgtgtctcttcttataagtaacagcgtctctgtgacctg 781
 |||||||
 DB 918 tgtgttcccatctaatgtgtgtctcttcttataagtaacagcgtctctgtgacctg 977
 QY 782 caaccaccttcgcccagcttccaagagcacaagaaagtgttctctctctatgacatc 841
 |||||||
 DB 978 caaccaccttcgcccagcttccaagagcacaagaaagtgttctctctctatgacatc 1037
 QY 842 attgcacacatgtcgaatcccttataatacaacttgagacaagaggttaagaagagc 901
 |||||||
 DB 1038 attgcacacatgtcgaatcccttataatacaacttgagacaagaggttaagaagagc 1097
 QY 902 tttaaaaggttggttcgaagagcttcttcaatcaagaataa 943
 |||||||
 DB 1098 tttaaaaggttggttcgaagagcttcttcaatcaagaataa 1139

RESULT 6
 US-09-800-321A-5

Sequence 5, Application US/09800321A
 GENERAL INFORMATION:
 APPLICANT: Padigaru, Muralidhara
 APPLICANT: Burgess, Catherine E
 APPLICANT: Mishra, Vishnu
 APPLICANT: Li, Li
 APPLICANT: Baumgartner, Jason C
 APPLICANT: Majumder, Kumud
 APPLICANT: Spytsek, Kimberly A
 APPLICANT: Tchernev, Velizar T
 TITLE OF INVENTION: Novel Proteins and Nucleic Acids Encoding Same
 FILE REFERENCE: 15966-703 US
 CURRENT APPLICATION NUMBER: US/09/800,321A
 PRIOR FILING DATE: 2001-03-05
 PRIOR APPLICATION NUMBER: 60/186,606
 PRIOR FILING DATE: 2000-03-03
 PRIOR APPLICATION NUMBER: 60/221,942
 PRIOR FILING DATE: 2000-07-31
 PRIOR APPLICATION NUMBER: 60/260,285
 PRIOR FILING DATE: 2001-01-08
 PRIOR APPLICATION NUMBER: 60/220,263
 PRIOR FILING DATE: 2000-07-24
 PRIOR APPLICATION NUMBER: 60/257,600
 PRIOR FILING DATE: 2000-12-21
 PRIOR APPLICATION NUMBER: 60/187,295
 PRIOR FILING DATE: 2000-03-06
 PRIOR APPLICATION NUMBER: 60/219,854
 PRIOR FILING DATE: 2000-07-21
 PRIOR APPLICATION NUMBER: 60/187,249
 PRIOR FILING DATE: 2000-03-06
 PRIOR APPLICATION NUMBER: 60/187,247
 PRIOR FILING DATE: 2000-03-06
 PRIOR APPLICATION NUMBER: 60/187,250
 PRIOR FILING DATE: 2000-03-06

```

: PRIOR APPLICATION NUMBER: 60/187,253
: PRIOR FILING DATE: 2000-03-06
: PRIOR APPLICATION NUMBER: 60/187,248
: PRIOR FILING DATE: 2000-03-06
: PRIOR APPLICATION NUMBER: 60/187,296
: PRIOR FILING DATE: 2000-03-06
: PRIOR APPLICATION NUMBER: 60/187,563
: PRIOR FILING DATE: 2000-03-07
: NUMBER OF SEQ ID NOS: 78
: SOFTWARE: Patentin Ver. 2.1
: SEQ ID NO 5
: LENGTH: 956
: TYPE: DNA
: ORGANISM: Homo sapiens
: US-09-800-321A-5

```

Query Match	99.48;	Score 937.2;	DB 31;	Length 956;
-------------	--------	--------------	--------	-------------

QY	2	atgaattcggtaaatgaaagacacatacaagaggttattctcgtcgtgttcctccacaatga	61
Db	7	atgaattcggtaaatgaaagacacatacaagaggttattctcgtcgtgttcctccaatgaa	66
QY	62	ccttcgtcgtgagtttcacactccctgtgtgtctctctgtattctcttaacactgtgacacatctt	121
Db	67	ccttcgtcgtgagtttcacactccctgtgtgtctctctgtattctcttaacactgtgacacatctt	126
QY	122	ggcaatcgcgaacattattcttagtgcacagcctcggagacacaaacttcataccccatgtat	181
Db	127	ggcaatcgcgaacattattcttagtgcacagcctcggagacacaaacttcataccccatgtat	186
QY	182	tttttcttaacaatctatacactcctctgtatctctgttaacaacaatgtacagttccaca	241
Db	187	tttttcttaacaatctatacactcctctgtatctctgttaacaacaatgtacagttccaca	246
QY	242	atgcgtagtaaatlttatgcagacatacgaagaagtaatcagttatcgtcgtcgtgtgacccag	301
Db	247	atgcgtagtaaatlttatgcagacatacgaagaagtaatcagttatcgtcgtcgtgtgacccag	306
QY	302	ctttccataattcttgccttggggtcctactgaatatctcttcctcgtccgtatgtccctt	361
Db	307	ctttccataattcttgccttggggtcctactgaatatctcttcctcgtccgtatgtccctt	366
QY	362	gattggtttgaagcattctgttcgcgcctccatactcaatgatcatatgacacagaagatc	421
Db	367	gattggtttgaagcattctgttcgcgcctccatactcaatgatcatatgacacagaagatc	426
QY	422	tgacctcagatttgcagcgtcgtacatcctctgggtttaaagtctttagtaactcagttgtgttct	481
Db	427	tgacctcagatttgcagcgtcgtacatcctctgggtttaaagtctttagtaactcagttgtgttct	486
QY	482	accctgaactctccagcgtcgtacatcctcgtgaacccctatgtgataagatcacacttctctgttaa	541
Db	487	accctgaactctccagcgtcgtacatcctcgtgaacccctatgtgataagatcacacttctctgttaa	546
QY	542	gtccctgcagtcgtccaaagtattatctgttctgttgagacaaagcaaatggaagctgaacatctc	601
Db	547	gtccctgcagtcgtccaaagtattatctgttctgttgagacaaagcaaatggaagctgaacatctc	606
QY	602	ctctgtcagtgaagctcttcacatactaaataccctcgtgacactcaatccattatcatatagctt	661
Db	607	ctctgtcagtgaagctcttcacatactaaataccctcgtgacactcaatccattatcatatagctt	666
QY	662	attgtcgcagcagtatcttgaagatatagtcctcgtcgtgaaggtcgacaaaagcaatttgggaca	721
Db	667	attgtcgcagcagtatcttgaagatatagtcctcgtcgtgaaggtcgacaaaagcaatttgggaca	726
QY	722	tttgtgtcccaatcgaattgtgtgtgtctctttttttagtaacagcgcgtctctcttgaacctg	781
Db	727	tttgtgtcccaatcgaattgtgtgtgtctctctttttttagtaacagcgcgtctctcttgaacctg	786

Qy	782	caacacacttcgcgcagctcccaagagacaaagaaagtgtttctctctctctatgtgaatc	84.1
Db	787	caacacacttcgcgcagctcccaagagacaaagaaagtgtttctctctctctatgtgaatc	84.6
Qy	842	attgcacccatgcgcgaatccctctatatatacaactttgagaacaagaggttaaaagaaagc	90.1
Db	847	attgcacccatgcgcgaatccctctatatatacaacttcgaaacgagaggttaaaagaaagc	90.6
Qy	902	tttaaaaggttggttcgaagaagctctctcttaatacaagaataaa	94.3
Db	907	tttaaaaggttggttcgaagaagctctctcttaatacaagaataaa	94.8

RESULT 7
US-60-261-976-77
; Sequence 77, Application US/60261976

```

1  APPLICANT: Beasley, Ellen
2  TITLE OF INVENTION: ISOLATED HUMAN G-PROTEIN COUPLED
3  TITLE OF INVENTION: RECEPTORS, NUCLEIC ACID MOLECULES ENCODING HUMAN GPCR
4  TITLE OF INVENTION: PROTEINS, AND USES THEREOF
5  FILE REFERENCE: CL001089
6  CURRENT APPLICATION NUMBER: US/60/261,976
7  CURRENT FILING DATE: 2001-01-17
8  NUMBER OF SEQ ID NOS: 465
9  SOFTWARE: FastSeq for Windows Version 4.0
10 SEQ ID NO 77
11 LENGTH: 20410
12 TYPE: DNA
13 ORGANISM: HUMAN
14 OS-60-261-976-77

```

Query Match	76.0%;	Score 716.8;	DB 65;	Length 20410;
Best Local Similarity	86.3%;	Pred. No. 1.5e-204;		
Matches 815;	Conservative 0;	Mismatches 127;	Indels 2;	Gaps 2

[illegible]


```

Oy      541 agccccgaactgctcaagttatcttgttgttagacaacacagcaaatgagcgtgaactt    600
Db      2191 agtcctctctctgcgtccaagtgtccctgtgttgacacacaaagaattggcgtaactt    2250
Oy      601 ccttgcagtagagctctcatcattaccacctgacactacctaactatacatagctt    660
Db      2251 cttaatacagtgctatctccctcattaacaccogtgaactcatcattatgtaagctt    2310
Oy      661 tatgtccgagcagtatattgagagatacagtcgtcgtgaaggctgcacaaaagaattggac    720
Db      2311 tatgtccaaagcagtggttgaaatccagtcgtcgtlaaaggtaacgaagaagcaattggac    2370
Oy      721 atgtgttcccacatcaattgtgtgtgtgtctctttttaatagcacagcgctcgtgtacct    780
Db      2371 atgtgtccccaatcatcaattgtgtgtgtgtctctttttaatagcacatcctcatgtacct    2430
Oy      781 gcaaccaccttcgccccagcgtcccaagaccaaagaagaatggttctctctctatgtaat    840
Db      2431 gcaaccaccttccacccagctcccaagaacccgggaagaatggttctctctctctgtgaat    2490
Oy      841 catgcaccacatgtcgaatccctctatatatacaatttggacaaggggttaagaagag    900
Db      2491 catgcaccacatgtcgaatccctctatatatacaactttagaacaagaagggttaagaagc    2550
Oy      901 cttaaagaagttgtgtgc-aagagctctcttaatcaagaanaaa    943
Db      2551 cttaaagaagttgtgtcaagaagctctcttaatcaagaanaaa    2594

RESULT:      8
US-60-360-207-43515
; Sequence 43515, Application US/60360207
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig
; TITLE OF INVENTION: HUMAN GENOME DISCOVERY SYSTEM AND USES THEREOF
; FILE REFERENCE: CLO001321
; CURRENT APPLICATION NUMBER: US/60/360,207
; CURRENT FILING DATE: 2002-03-01
; NUMBER OF SEQ ID NOS: 47235
; SEQ ID NO 43515
; LENGTH: 1047
; TYPE: DNA
; ORGANISM: HUMAN
US-60-360-207-43515
```

Query Match	70.5%;	Score 665.2;	DB 75;	Length 1047;
Best Local Similarity	81.6%;	Pred. No. 1,4e-189;		
Matches 769;	Conservative	0;	Mismatches 173;	Indels 0;
				Gaps 0;
0Y	2	atgaattggtgaagaagacagcatcatcacagagattatctgcgtggtttctccaatgca	61	
Db	106	atgagcttggcgcaaacagagacatcacagcgaggtttgttctctcttgagttcttcatacaa	165	
0Y	62	ccttgctggagattccagatcccttctgtgtctctgattctcttaacagtcagcaatcctt	121	
Db	166	ccatgctgcggaggttcccaactcttctgtgtctctctgacatctctaagtgtgacacatctt	225	
0Y	122	ggcaactctgacatcatattctagtgtaacgcttgagaccacaaatctacatccccatgtat	181	
Db	226	ggaatctccacaactattctctgtgtcccatgttgatctctaactccacaatcccatgttac	285	
0Y	182	ttttcttacaacatctatcaactccctgtgattcttgtttacaaccaatgtacagttccacaa	241	
Db	286	tctctcttgactaaatctgtcagtcataagatcttcttcatatcaacatgacagttccacaa	345	
0Y	242	atgcagtagtaaatatttgagcatcagcgaggaagaatcagttatcagcgtgctgttagccag	301	
Db	346	atgcctggtaaacattctgcagcatcttgaggaagtaaaagtcttctggcgtctgtgtgtccag	405	
0Y	302	ctttcataattctcggccttggggactactgtataatctctctctcgtgcgcgtcatgtccctt	361	
Db	406	ctttcataattctcggccttggggactactgtataatctctctctcgtgcgcgtcatgtccctt	465	

QY	362	gacgtgttgtagctatcttctggtccctccatctactcaagatcatcaatgacacagagctc	421
Db	466	gacaggttttagagccataatctgcgcctccatctactcaagatcatcaatgacacagagctc	525
QY	422	tgccctcaagttggagagctgatcgtacgttggttactggttttagtaactcaagttggtgtc	481
Db	526	tgccctcaagttggagagctgatcgtatccctggatcaatttggcttggaaactcgggtggtcatcc	585
QY	482	aacctcaactctccagcgcgcacactctctgacacccctatgtagatgaatcctctcgttgaa	541
Db	586	attctaaactctccagcgttgccaacgctgtggccactatgctatagacacactctctgttgaa	645
QY	542	gtccctcactgtctcaagttatctctgtgttgagacaacagcaaatgtaggctgaaactatcc	601
Db	646	gtgcctcagcagctccaagttgtctctgtgttagtgaacagacaatgaagctgaactgtctc	705
QY	602	ctgtgtcagttgagcctctccatctcaataacccctgacactcatctctatcatatgctttt	661
Db	706	cttggagagttgtctctccacccaacccctctctccatccatccatcaacataatgacatt	765
QY	662	atgttccgagcagatgttagagatatacagctctgcgcgaagctgcgacaacaaagcatcttgagaca	721
Db	766	atagcccggtgcacatcttgaanaatccaatccatctgcgcgaaggtcgtccaanaagcgttttgagacc	825
QY	722	tgtgtgtcccatctaatgttggtgtgtctctttttatagtagacagccgtctctgtgtacctg	781
Db	826	tgtagctctccatctcaataagtagtgaatacactgttttaagtgaacagctctctccgttgaactt	885
QY	782	caacacacttggcccgactctcaagagacacaaggaagaatggttctctctctcatgaaatc	841
Db	886	ctaacacacatcaacacacatccccaagaagaagaagatggtgtccacttttctcatgaaatc	945
QY	842	attgcacacatgcgtgaatctccctatataataccttaggaacaagaaggtlaaaggaagc	901
Db	946	attgcccccatgcgtgaaccccttatctataacacttaggaacaagaaggtlaaaggaagc	1005
QY	902	tttaaaaggtgtgttcaagaagctctcttcaatcaagaanaaa	943
Db	1006	tttaaaaggtgtgacaaagagagctctctttagaacaanaactaa	1047

```

RESULT 9
US-60-360-207-42178
; Sequence 42178, Application US/60360207
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig
; TITLE OF INVENTION: HUMAN GENOME DISCOVERY SYSTEM AND USES THEREOF
; FILE REFERENCE: CLO01321
; CURRENT APPLICATION NUMBER: US/60/360, 207
; CURRENT FILING DATE: 2002-03-01
; NUMBER OF SEQ ID NOS: 47235
; SEQ ID NO 42178
; LENGTH: 1071
; TYPE: DNA
; ORGANISM: HUMAN
US-60-360-207-42178

Query Match          63.9%  Score 602.2;  DB 75;  Length 1071;
Best Local Similarity 77.4%;  Pred. No. 1.6e-170;
Matches 730;  Conservative 0;  Mismatches 215;  Indels 0;  Gaps 0.

QY 1 tatgaattggtgaatagacagacatatacagaagattatccgtcgtgtgttcagatcg 60
    |||||  |||  |||||  |||||  |||||  |||||  |||  |||||
DB 129 tatggtgtggtgacatgagatcatctcaacggaagttaattcctcttaggttttcagatg 188
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 61 accttgctgtgagtttcaactcctctgtgtgtctctcttgattcttaacctgtgacatctt 120
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 189 gccatgtgtgagctgtgcgcctctcttgtgtgtgttcttagtgccttatctgacacatctt 248
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 121 ttgcaatcgtcacattatctagtggtcaacgctgtgacacacaaacttcatacccatgta 180
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

```

```
Db 249 tggaaataatgatcatcttctgtgtccgcctgagatccaaactcacaaccccatgta 308
Qy 181 tttttcttaccatcatcatccctcgtgactcttctgttaacacacatgtaacgccaca 240
    ||||| || || || || || || || || || || || || || || || || || || || ||
Db 309 cttttcttaccatcatccctcgtgactcttctgttaacacacacatgtaacgccaca 368
Qy 241 aatgctagtaattatgacacatcaggaagtaatacagtaacgtgctgtgttagcca 300
    ||||| || || || || || || || || || || || || || || || || || || || ||
Db 369 gatgcatcaacaacatcgcagcaccgggaagtgatcagctatgtgtgtgtgtgtgtgt 428
Qy 301 gctttcaatatttcgctccttgggggagtaataatcctctcctcctcctcctcctcct 360
    ||||| || || || || || || || || || || || || || || || || || || || ||
Db 429 gctttcaatatttcgctccttgggggagtaataatcctctcctcctcctcctcctcct 488
Qy 361 tgattgttctgttagctattctgtcgcctcctcctcctcctcctcctcctcctcctcct 420
    ||||| || || || || || || || || || || || || || || || || || || || ||
Db 489 tgaacagttgttagcactcgtcgcctcctcctcctcctcctcctcctcctcctcctcct 548
Qy 421 ctgctcctcagttgagcagctcagctcctcctcctcctcctcctcctcctcctcctcct 480
    ||||| || || || || || || || || || || || || || || || || || || || ||
Db 549 ctgctcctcagttgagcagctcagctcctcctcctcctcctcctcctcctcctcctcct 608
Qy 481 taacctgacctcagctcagctcagctcctcctcctcctcctcctcctcctcctcctcctc 540
    ||||| || || || || || || || || || || || || || || || || || || || ||
Db 609 taagtgaacctcagctcagctcagctcctcctcctcctcctcctcctcctcctcctcctc 668
Qy 541 agtccctcagctcagctcagctcagctcctcctcctcctcctcctcctcctcctcctcct 600
    ||||| || || || || || || || || || || || || || || || || || || || ||
Db 669 agtccctcagctcagctcagctcagctcctcctcctcctcctcctcctcctcctcctcct 728
Qy 601 cctgtttagtgagctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcct 660
    ||||| || || || || || || || || || || || || || || || || || || || ||
Db 729 ctcatcaatgctgtgttcttcttaataccggtgacccctcctcctcctcctcctcctcct 788
Qy 661 tatgtcgcagcagctatctgagatcagctcgtcgtgagacagacaaatgagctgtgagac 720
    ||||| || || || || || || || || || || || || || || || || || || || ||
Db 789 tatgtcgcagcagctatctgagatcagctcgtcgtgagacagacagctgtgagacagct 848
Qy 721 atgtgtgtcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcct 780
    ||||| || || || || || || || || || || || || || || || || || || || ||
Db 849 atgtgtgtcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcct 908
Qy 781 gcaaccctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcct 840
    ||||| || || || || || || || || || || || || || || || || || || || ||
Db 909 gcaaccctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcct 968
Qy 841 catgcacacacatgctgtaacccctcctcctcctcctcctcctcctcctcctcctcctc 900
    ||||| || || || || || || || || || || || || || || || || || || || ||
Db 969 catgcacacacatgctgtaacccctcctcctcctcctcctcctcctcctcctcctcct 1028
Qy 901 cttaaaaggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 943
    ||||| || || || || || || || || || || || || || || || || || || || ||
Db 1029 gtccaagaggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1071
```

```
RESULT 10
US-60-178-308-604/c
; Sequence 604, Application US/60178308
; GENERAL INFORMATION:
; APPLICANT: Bonazzi, Vivien
; TITLE OF INVENTION: ISOLATED HUMAN G-PROTEIN COUPLED
; TITLE OF INVENTION: RECEPTORS, NUCLEIC ACID MOLECULES ENCODING HUMAN GPCR
; FILE REFERENCE: C1000204
; CURRENT APPLICATION NUMBER: US/60/178,308
; CURRENT FILING DATE: 2000-01-27
; NUMBER OF SEQ ID NOS: 3344
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 604
; LENGTH: 673
; TYPE: DNA
; ORGANISM: HUMAN
US-60-178-308-604
```

```
Query Match 63.7%; Score 600.4; DB 56; Length 673;
Best Local Similarity 99.8%; Pred. No. 4.5e-170;
Matches 601; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 342 tccctgacgctatgtctccttctgtatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 401
    ||||| || || || || || || || || || || || || || || || || || || || ||
Db 673 TCCGAGCCGATCATCTCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 614
Qy 402 ttatcagacacacacacacacacacacacacacacacacacacacacacacacacacacacac 461
    ||||| || || || || || || || || || || || || || || || || || || || ||
Db 613 TTATCATGACACACACACACACACACACACACACACACACACACACACACACACACACACACAC 554
Qy 462 gtaactagctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 521
    ||||| || || || || || || || || || || || || || || || || || || || ||
Db 553 GTAACTAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 494
Qy 522 taagtaacttctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcct 581
    ||||| || || || || || || || || || || || || || || || || || || || ||
Db 493 TAGATCATCTTCTCTGTGAAGTCCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTC 434
Qy 582 caaatgagctgtgactatctcctcctcctcctcctcctcctcctcctcctcctcctcctcct 641
    ||||| || || || || || || || || || || || || || || || || || || || ||
Db 433 CAATGTAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 374
Qy 642 tccctatataatgcttcttctcctcctcctcctcctcctcctcctcctcctcctcctcctc 701
    ||||| || || || || || || || || || || || || || || || || || || || ||
Db 373 TCCATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 314
Qy 702 gacaaaagacatctgtgagatctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 761
    ||||| || || || || || || || || || || || || || || || || || || || ||
Db 313 GACAAAAAGCATTTGGGACATGTGGTCCCATCATTAATGTGTGTGTGTGTGTGTGTGTGT 254
Qy 762 cagcgcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcct 821
    ||||| || || || || || || || || || || || || || || || || || || || ||
Db 253 CAGCGCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 194
Qy 822 ttctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcct 881
    ||||| || || || || || || || || || || || || || || || || || || || ||
Db 193 TTCTCTCTCTCTATGAAATGATGACCCATGCAATGCAATGCAATGCAATGCAATGCAATG 134
Qy 882 acaagaggtgaagaggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 941
    ||||| || || || || || || || || || || || || || || || || || || || ||
Db 133 ACAAGGAGGTGAAGGAGGTGAAGGAGGTGAAGGAGGTGAAGGAGGTGAAGGAGGTGA 74
Qy 942 aa 943
    ||
Db 73 AA 72
```

```
RESULT 11
US-60-245-225-86/c
; Sequence 86, Application US/60245225
; GENERAL INFORMATION:
; APPLICANT: Beasley, Ellen
; TITLE OF INVENTION: ISOLATED HUMAN G-PROTEIN COUPLED
; TITLE OF INVENTION: RECEPTORS, NUCLEIC ACID MOLECULES ENCODING HUMAN GPCR
; FILE REFERENCE: C1000885
; CURRENT APPLICATION NUMBER: US/60/245,225
; CURRENT FILING DATE: 2000-11-03
; NUMBER OF SEQ ID NOS: 705
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 86
; LENGTH: 46924
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)...(46924)
; OTHER INFORMATION: n = A,T,C or G
US-60-245-225-86
```

Query Match 56.6%; Score 534; DB 63; Length 46924;
Best Local Similarity 72.9%; Pred. No. 4.1e-149;
Matches 687; Conservative 0; Mismatches 255; Indels 0; Gaps 0;

QY 2 atgaattggtaaatgacgaacatcatatcagagatttcttgcgtgggttcttcagatcga 61
 DB 45060 ATGAATTGGGAAAATGAGAGCTCCCAAAAAGATTATACTACTTGGCTTCGAAAGG 45001
 QY 62 ccttgagctgagttccaccccttggctctcttgattcttcttcaacgttgagacattt 121
 DB 45000 GCTTGGCTAAATAATGCCCCCTTTTGTGTGCTCTGTTAATATCATACAAATCACCATTAT 44941
 QY 122 ggaacatcgacatattatctagtgacagcctggacacaaacttcaaacccatglat 181
 DB 44940 GGCATATGTCTCATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 44881
 QY 182 ttttttcttcaacatctatcactctctgattcttcttgaacacacatgtacagttccaca 241
 DB 44880 TTCTTTCTCACTAATCTCTCACTAATCTCTCACTAATCTCTCACTAATCTCTCACTAATCTCT 44821
 QY 242 atgctagtaaatattatcgacatcagaagaatactagttatctgctgtgtatgagccag 301
 DB 44820 ATGCTTAGTAATATTGGTTGCAACAAAAGACATCACTATGCTGTGCTGTGCTGTGCTGTGCTGTG 44761
 QY 302 ctcttcaatattctggccttggggtgctactgaatattcttctcctggcgtatgctctt 361
 DB 44760 CTCAATCATCTCTCTGGCCCTAGGTGCTACAGAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 44701
 QY 362 gattggtttgagctatttgcgcctctcactatcactcaatgattacagacagagactc 421
 DB 44700 GACAGATATGTGGCTGTTGTCAGACCCCTCCACATATGATGATGATGATGATGATGATGATGATG 44641
 QY 422 tgcctcagttggagcagctgacatcagctgagtttgaatttgaattgagttgttctt 481
 DB 44640 TGCTTAAGATGAGGAGGCTTCTCATGGCTCATTTGGTTTGGCAACTCATAGTCTCATGCTCT 44581
 QY 482 accctgactctcagcagctgacactctgtgacccctatgtatagatcaacttctctgtgaa 541
 DB 44580 TCCTTGAAGTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 44521
 QY 542 gtccctcagctgctcgaattatcttctgtgtgagacacagaaatgtgagctgaactatc 601
 DB 44520 GTGCTGCTCACTCTCAAGATTGTCATGTCGTCACAAAGCCCTATTTGAGGCTCTTC 44461
 QY 602 ctgtcagtgagcctctcactatcactaccctgacactcactatatacatatgacttt 661
 DB 44460 TTCTTTAGTGTACTAATTTCTTAATTCAGATGATGATGATGATGATGATGATGATGATGATGAT 44401
 QY 662 atgtccgagcagatattgagagatacagctgtgagagtgacagacaaagaactttgggaca 721
 DB 44400 ATAGCTCAACAGATTAATAAATCAGATGATGATGATGATGATGATGATGATGATGATGATGAT 44341
 QY 722 tgttggtccatcaatattgt 781
 DB 44340 TGTGTTGTCACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 44281
 QY 782 caaacactctcagcagctcgaagcgaagaagaatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 841
 DB 44280 CAACACACTTCAATCCACTCTCAAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 44221
 QY 842 atgtcaccatgtcgtgaatccccctatatatacacttgaagacaaagaagaatgaagac 901
 DB 44220 ATCAATATCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 44161
 QY 902 tttaaaagttgtgtgagagagtttcttcaatcaagaataa 943
 DB 44160 TTCAAGAGGCTGATGCCAAGATCTTTTCTGTAAAGAAATA 44119

; Sequence 27, Application US/60258250
 ; GENERAL INFORMATION:
 ; APPLICANT: Beasley, Ellen
 ; TITLE OF INVENTION: ISOLATED HUMAN G-PROTEIN COUPLED
 ; TITLE OF INVENTION: RECEPTORS, NUCLEIC ACID MOLECULES ENCODING HUMAN GPCR
 ; TITLE OF INVENTION: PROTEINS, AND USES THEREOF
 ; FILE REFERENCE: C1001035-PROV
 ; CURRENT APPLICATION NUMBER: US/60/258,250
 ; CURRENT FILING DATE: 2000-12-27
 ; NUMBER OF SEQ ID NOS: 210
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 27
 ; LENGTH: 46924
 ; TYPE: DNA
 ; ORGANISM: HUMAN
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (1)..(46924)
 ; OTHER INFORMATION: n = A,T,C or G
 US-60-258-250-27

Query Match 56.6%; Score 534; DB 64; Length 46924;
Best Local Similarity 72.9%; Pred. No. 4.1e-149;
Matches 687; Conservative 0; Mismatches 255; Indels 0; Gaps 0;

QY 2 atgaattggtaaatgacgaacatcatatcagagatttcttgcgtgggttcttcagatcga 61
 DB 45060 ATGAATTGGGAAAATGAGAGCTCCCAAAAAGATTATACTACTTGGCTTCGAAAGG 45001
 QY 62 ccttgagctgagttccaccccttggctctcttgattcttcttcaacgttgagacattt 121
 DB 45000 GCTTGGCTAAATAATGCCCCCTTTTGTGTGCTCTGTTAATATCATACAAATCACCATTAT 44941
 QY 122 ggaacatcgacatattatctagtgacagcctggacacaaacttcaaacccatglat 181
 DB 44940 GGCATATGTCTCATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 44881
 QY 182 ttttttcttcaacatctatcactctctgattcttcttgaacacacatgtacagttccaca 241
 DB 44880 TTCTTTCTCACTAATCTCTCACTAATCTCTCACTAATCTCTCACTAATCTCTCACTAATCTCT 44821
 QY 242 atgctagtaaatattatcgacatcagaagaatactagttatctgctgtgtatgagccag 301
 DB 44820 ATGCTTAGTAATATTGGTTGCAACAAAAGACATCACTATGCTGTGCTGTGCTGTGCTGTGCTGTG 44761
 QY 302 ctcttcaatattctggccttggggtgctactgaatattcttctcctggcgtatgctctt 361
 DB 44760 CTCAATCATCTCTCTGGCCCTAGGTGCTACAGAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 44701
 QY 362 gattggtttgagctatttgcgcctctcactatcactcaatgattacagacagagactc 421
 DB 44700 GACAGATATGTGGCTGTTGTCAGACCCCTCCACATATGATGATGATGATGATGATGATGATGATG 44641
 QY 422 tgcctcagttggagcagctgacatcagctgagtttgaatttgaattgagttgttctt 481
 DB 44640 TGCTTAAGATGAGGAGGCTTCTCATGGCTCATTTGGTTTGGCAACTCATAGTCTCATGCTCT 44581
 QY 482 accctgactctcagcagctgacactctgtgacccctatgtatagatcaacttctctgtgaa 541
 DB 44580 TCCTTGAAGTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 44521
 QY 542 gtccctcagcagctcgaattatcttctgtgtgagacacagaaatgtgagctgaactatc 601
 DB 44520 GTGCTGCTCACTCTCAAGATTGTCATGTCGTCACAAAGCCCTATTTGAGGCTCTTC 44461
 QY 602 ctgtcagtgagcctctcactatcactaccctgacactcactatatacatatgacttt 661
 DB 44460 TTCTTTAGTGTACTAATTTCTTAATTCAGATGATGATGATGATGATGATGATGATGATGATGATGAT 44401
 QY 662 atgtccgagcagatattgagagatacagctgtgagagtgacagacaaagaactttgggaca 721
 DB 44400 ATAGCTCAACAGATTAATAAATCAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 44341

```
QY 722 tctgtcccaatcaatgtgtgtctcttttataagtaagccgtctctgttacctg 781
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 44340 TGTGGTCCACATGATGTGTGTGTCCTTTATATGGAACGACCATTTATATGATCTT 44281

QY 782 caaccaccctccagcgtcccaaggaagaagatgtgtctctctctatgtgaac 841
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 44280 CAACCCACTTCATCCACTCTTAAGGACTGGGAAAGATGGTTCCCTCTCTATGGAATC 44221

QY 842 atgcaccatgtcgaatccctctataataactgaaggaagaaggaaggaagc 901
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 44220 ATCAATCCATGTGGAAGTCCCTCATCTACAGCCTTAGAAATTAAGATATGAAGAGGCC 44161

QY 902 tttaaaagtgtgtgaagagcttcttaacaagaataa 943
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 44160 TTCAAGAGCGCTGATGCCAAGATCTTTTCTGTGAAGAAATTA 44119

RESULT 13
US-60-248-505-274/c
; Sequence 274, Application US/60248505
; GENERAL INFORMATION:
; APPLICANT: Beasley, Ellen
; TITLE OF INVENTION: ISOLATED HUMAN G-PROTEIN COUPLED
; TITLE OF INVENTION: RECEPTORS, NUCLEIC ACID MOLECULES ENCODING HUMAN GPCR
; FILE REFERENCE: c1000918
; CURRENT APPLICATION NUMBER: US/60/248,505
; CURRENT FILING DATE: 2000-11-15
; NUMBER OF SEQ ID NOS: 1998
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 274
; LENGTH: 68686
; TYPE: DNA
; ORGANISM: human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(68686)
; OTHER INFORMATION: n = A,T,C or G
US-60-248-505-274

Query Match 56.6%; Score 534; DB 63; Length 68686;
Best Local Similarity 72.9%; Pred. No. 4.9e-149;
Matches 687; Conservative 0; Mismatches 255; Indels 0; Gaps 0;

QY 2 atgaattggtaaatgacagcatatcacaggaattatctgcgcgggttctcagatcga 61
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 66687 ATGAATTGGGAATAATGAGAGCTCCCAAAAGAGTTTACTACTTGGCTTCTCAGATAGG 66628

QY 62 ccttgctggaagttccactcctgtgtctctctgaattctctacactgtgacatctt 121
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 66627 GCTTGGCTAACAAATGCCCTTTTGTGTGCTGTTAATATCATATACAAATCAGCATATTT 66568

QY 122 ggcgaatcgaacatattcagtgtaagcgctggaacccaacttaaccccattat 181
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 66567 GGCATATGTTCATCATGATGAGTGTGATCTGTGATCCCAAACTTCATACCTCCATGTAT 66508

QY 182 ttttttttccaactatcatcactcctgactcttctgttaacacaagtlacagttccaca 241
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 66507 TTCTTTTCTCAATATCTCTCATCTTTAGATCTCTGCTATACCAACATACAGTCCCTCAT 66448

QY 242 atgctagtaaatatlatgacatcatcagaagaagtaactagttatcgtgtgtgtgacccag 301
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 66447 ATGTGTGTAATATATGTGTGCAACAATAAGCCATCAAGTATGCTGTGTGTGCCAC 66388

QY 302 ctttccatattctcgtccttgagggtgactgaataatctctcctgcgcgtcatgtctt 361
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 66387 CTCATCATCTTCCCTGGCCCTAGGTGATACAGAGTGTCTCTTCTGCTGTATATCTCTT 66328

QY 362 gattgtgttttagctatttctgagcctctcatattacagttatcatcagaccggaagctc 421
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 66327 GACAGATATGTGCTGTTTGACAGACCCCTCCACTATATGTAGTATCATATGAATTTATGTGTT 66268
```

```
QY 422 tgcctccagttgacgtcgtcatccctggtgttactgttttagtaactcagtggtgttct 481
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 66267 TGCTTAAGGATGGGACGCTTCTCATAGGCTCATTTGGTTTGGCAACCTCAGTCTCAGACT 66208

QY 482 accctgaactcagcgtcgcacactctgtgaaccctatgtatgatcatcttctgtgaa 541
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 66207 TCCCTTGTACTTAAACATAGCCCGCTGTGTCCACAGGAAGTGGACCACTTTTCTGTGAG 66148

QY 542 gtccctcactgctcgaattatctctgtgttgaagaacaagaagaatgaggtcgaactatc 601
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 66147 GTGCTTGACACTTTCAGAGTTTCTCATGTGCTGACACAAAGCCTATTGAGGCTGACTTTC 66088

QY 602 ctgtcagtgagctcttccatctaatccctcgtgacactcctctatcatcatgtcttt 661
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 66087 TTCTTATGTAGTAAATATCTTCTTAATTCAGATGACATGATCATCTCATCTATAGGCTTC 66028

QY 662 atgtccgagcagatattgagaatacagttctgtgaagtgacaaagaatltgggaca 721
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 66027 ATAGCTCAAGCAGATTAATAAATCAGGTCAACAGACAGACGCAAAAAGCATTTTGGGACA 65968

QY 722 tctgtcccaatcaatgtgtgtctcttcttataagtaagccgtctctgttacctg 781
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 65967 TGTGGTCCACATGATGTGTGTGTCCTCTTTATATGGAACGACCATTTATATGATCTT 65908

QY 782 caaccaccctccagcgtcccaaggaagaagatgtgtctctctctatgtgaac 841
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 65907 CAACCCACTTCATCCACTCTTAAGGACTGGGAAAGATGGTTCCCTCTCTATGGAATC 65848

QY 842 atgcaccatgtcgaatccctctataataactgaaggaagaaggaaggaagc 901
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 65847 ATCAATCCATGTGGAAGTCCCTCATCTACAGCCTTGAAGAAATTAAGATGAAGAGGCC 65788

QY 902 tttaaaagtgtgtgaagagcttcttaacaagaataa 943
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 65787 TTCAAGAGCGCTGATGCCAAGATCTTTTCTGTGAAGAAATTA 65746

RESULT 14
US-60-248-505-1606
; Sequence 1606, Application US/60248505
; GENERAL INFORMATION:
; APPLICANT: Beasley, Ellen
; TITLE OF INVENTION: ISOLATED HUMAN G-PROTEIN COUPLED
; TITLE OF INVENTION: RECEPTORS, NUCLEIC ACID MOLECULES ENCODING HUMAN GPCR
; FILE REFERENCE: c1000918
; CURRENT APPLICATION NUMBER: US/60/248,505
; CURRENT FILING DATE: 2000-11-15
; NUMBER OF SEQ ID NOS: 1998
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1606
; LENGTH: 1510
; TYPE: DNA
; ORGANISM: Human
US-60-248-505-1606

Query Match 56.1%; Score 529.4; DB 63; Length 1510;
Best Local Similarity 72.7%; Pred. No. 1.9e-148;
Matches 683; Conservative 0; Mismatches 256; Indels 0; Gaps 0;

QY 2 atgaattggtaaatgacagcatatcacaggaattatctgcgtgggttctcagatcga 61
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2 atgaattgggaataatgagagcttcccaaaagattatactactgtgtctctcagatag 61

QY 62 ccttgctggaagttccactcctgtgtctctctgaattctctacactgtgacatctt 121
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 62 gcttggtataaagtgcccttctgtgtctcgtgttaataatcatcaacaatcaacatatt 121

QY 122 ggcgaatcgaacatattcagtgtaagcgctggaacccaacttaaccccattat 181
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 122 ggcgaatgttccatcatgaltggtgtgacttctgtgaatcccaaaacttaactccatgtat 181
```


GenCore version 4.5
Copyright (c) 1993 - 2000 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 25, 2002, 17:54:57 ; Search time 432.84 Seconds
(without alignments)
5945.102 Million cell updates/sec

Title: US-09-800-321A-3

Perfect score: 943
Sequence: 1 tatgaattgggtaaatgaca.....tcttcttaataagaataaa 943

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1737567 seqs, 1364410299 residues

Total number of hits satisfying chosen parameters: 3475134

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_NA_New:*
1: /cgn2_6/ptodata/2/pna/PCT_NEW.COMB.seq:*
2: /cgn2_6/ptodata/2/pna/US06_NEW.COMB.seq:*
3: /cgn2_6/ptodata/2/pna/US07_NEW.COMB.seq:*
4: /cgn2_6/ptodata/2/pna/US08_NEW.COMB.seq:*
5: /cgn2_6/ptodata/2/pna/US09_NEW.COMB.seq:*
6: /cgn2_6/ptodata/2/pna/US10_NEW.COMB.seq:*
7: /cgn2_6/ptodata/2/pna/US60_NEW.COMB.seq:*
8: /cgn2_6/ptodata/2/pna/US60_NEW.COMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	448	47.5	2940	7	US-10-027-632-111618
2	403.6	42.8	967	7	US-10-032-189-33
3	402	42.6	939	7	US-10-032-106-2
4	371.6	39.4	1014	5	US-09-965-422-9
5	371.6	39.4	1187	1	PCT-US02-03635-59
6	370	39.2	1007	5	US-09-965-422-7
7	370	39.2	1013	5	US-09-965-422-3
8	370	39.2	1014	5	US-09-965-422-5
9	370	39.2	1065	5	US-09-965-422-7
10	357	37.9	1005	5	US-09-965-422-11
11	357	37.9	1124	1	PCT-US02-03635-80
12	345	36.6	1035	5	US-09-546-986A-1
13	343.2	36.4	2197	1	PCT-US02-03635-96
14	319.6	33.9	1351	5	US-09-546-986A-5
15	318	33.7	971	5	US-09-981-566A-27
16	318	33.7	1162	1	PCT-US02-09923-112
17	309.4	32.8	942	5	US-09-907-218-9
18	309.4	32.8	1351	1	PCT-US02-09923-106
19	297.8	31.6	975	6	US-10-023-601-57
20	291.4	30.9	1411	5	US-09-546-986A-3
21	289.4	30.7	1188	6	US-10-210-782-1
22	289	30.6	961	6	US-10-023-601-59
23	275.6	29.2	1098	7	US-10-149-826-65
24	273	29.0	1026	1	PCT-US02-03635-82
25	270.2	28.7	987	6	US-10-023-601-29

26	265	28.1	806	7	US-10-027-632-29717	Sequence 29717, A
27	262	27.8	973	1	PCT-US02-09923-116	Sequence 116, App
28	255.8	27.1	963	6	US-10-025-806-55	Sequence 55, App
29	255.6	27.1	945	6	US-10-098-754-21038	Sequence 21038, A
30	255.6	27.1	1301	1	PCT-US02-09923-91	Sequence 91, App
31	254.4	27.0	1134	1	PCT-US02-09923-133	Sequence 133, App
32	253.4	26.9	938	5	US-09-965-422-35	Sequence 35, App
33	253.4	26.9	1701	1	PCT-US02-09923-114	Sequence 114, App
34	252.6	26.8	2053	1	PCT-US02-09923-83	Sequence 83, App
35	251.6	26.7	948	5	US-09-965-422-23	Sequence 23, App
36	250.6	26.6	941	6	US-10-023-601-75	Sequence 75, App
37	250	26.5	948	5	US-09-965-422-19	Sequence 19, App
38	248.8	26.4	966	6	US-10-025-806-109	Sequence 109, App
39	248.8	26.4	966	6	US-10-025-806-111	Sequence 111, App
40	248.4	26.3	949	5	US-09-965-422-21	Sequence 21, App
41	248	26.3	1175	1	PCT-US02-01339-26	Sequence 26, App
42	247.4	26.2	941	6	US-10-023-601-73	Sequence 73, App
43	246.8	26.2	1023	6	US-10-023-601-37	Sequence 37, App
44	246.6	26.2	971	6	US-10-023-601-79	Sequence 79, App
45	245.8	26.1	1451	1	PCT-US02-03635-67	Sequence 67, App

ALIGNMENTS

```
RESULT 1
US-10-027-632-111618
; Sequence 111618, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 111618
; LENGTH: 2940
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-111618

Query Match 47.5% Score 448; DB 7; Length 2940;
Best Local Similarity 73.2% Pred. No. 7.6e-116;
Matches 574; Conservative 0; Mismatches 210; Indels 0; Gaps 0;

OY 160 caaatcatatccccatgtatcttttcttcaaatcatatcatctctgtatcttcttca 219
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 caaatcatatccccatgtatcttttcttcaaatcatatcatctctgtatcttcttca 60

OY 220 caccacatgtacagtcacacaaatgtcagtaattatgcagcatcgggaagtaatcag 279
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 61 taccacacactacagtcacacaaatgtcagtaattatgcagcatcgggaagtaatcag 120

OY 280 ttatcgtgagcgtgttagccagcgtttcatatcttgcgtcgttggggaactgaatct 339
    ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 121 ctatcgtgagcgtgttagccacccatcatctcttcgtgacctaggtgctaacagagtct 180
```

Qy	340	ttcttcggccctaatgctccctttgtttgtttgttagcattttgtcgcctcccaattc	339
Db	181	cccttcgtcgtttatgtcttcctttgcacagatagtgtcgtttgttcagaccctccctc	240
Qy	400	agttatcatgtcaccaagagactctgcctccagttgttcagcgttcacctcgtgtta	459
Db	241	agttacatgatattatgttcttcgtcctaagatgtgcagccctctccatggtcatt	300
Qy	460	tagttacccaattgtgtgtgttaccctgcacctcccaagctcgaactctgtgaacct	519
Db	301	cggcaaccacagttgcgtcaagctcttcctttacactcttaacatgccaagcgtgtg	360
Qy	520	gatatatcattctctctgttagagtccttcgcacgtcgtccaagtatccttgtttg	579
Db	361	agttgacacactttctctgttgaggtgtgcctgcacctctccaagttgtcattg	420
Qy	580	agcaaatgaaggctgaactatctccttgcattgaagctctccatccataaccctg	639
Db	421	gacctattgaggtcgtgcctctctctctttagtgtagtaactcttccatactcc	480
Qy	640	catccttatcatatagcttttatgtccgagccagttatgttaggtataagctgt	699
Db	481	gacctctcctccctatgtccttcattagctccaagcagttatlaaaatcaggt	540
Qy	700	tcgaacaaaacatttggagacatgtgttcccatctaatgtgtgtcctctttta	759
Db	541	acggcaaaaagcatttggagaaatgtgtgtgtcccaatagtattgtgtcctct	600
Qy	760	tacagccgtctctgttgaacctgcgaacacacttcggccagctcccaaggaca	819
Db	601	aaccgacctattatgttatctctcaaccacacttccatccacttaagacctg	660
Qy	820	ggttcctctctcattgaaatcatcttgcaaccatctgaatcccttatatac	879
Db	661	ggttccctctctcatggaacacacacatccatgattgaactcctctccacag	720
Qy	880	gaacaaaggaggtaaaaggagcctttaaaggttgtgtgcagaagctctctta	939
Db	721	aaataagatatgaggagggcctccaagagctgatgtccaagaatcctttctg	780
Qy	940	ataa 943	
Db	781	ataa 784	
RESULT 2			
US-10-032-189--33			
Sequence 33, Application US/10032189			
GENERAL INFORMATION:			
APPLICANT: Alsobrook II, John P			
APPLICANT: Tcherniev, Velizar T			
APPLICANT: Liu, Xiaohong			
APPLICANT: Szytek, Kimberly A			
APPLICANT: Zernusen, Bryan D			
APPLICANT: Patturajan, Meera			
APPLICANT: Grose, William M			
APPLICANT: Lepley, Denise M			
APPLICANT: Burgess, Catherine E			
APPLICANT: Shumets, Richard A			
APPLICANT: Grose, William M			
APPLICANT: Szekeres, Edward S			
APPLICANT: Vernet, Corine A.M.			
APPLICANT: Li, Li			
APPLICANT: Casman, Stacie J			
APPLICANT: Boldog, Ferenc L			
APPLICANT: Gorman, Linda			
APPLICANT: Gangoli, Esha A			
APPLICANT: Fernandes, Elma R			
APPLICANT: Rieger, Daniel K			
APPLICANT: Edinger, Shlomit R			
APPLICANT: Gunther, Erik			
APPLICANT: Millet, Isabelle			

	APPLICANT:	Scioren, Paul
	APPLICANT:	Ellieman, Karen
	APPLICANT:	MacDougall, John R
	APPLICANT:	Smithson, Glenda
	TITLE OF INVENTION:	Proteins and Nucleic Acids Encoding Same
	FILE REFERENCE:	21402-228
	CURRENT APPLICATION NUMBER:	US/10/032.189
	CURRENT FILING DATE:	2001-12-21
	PRIOR APPLICATION NUMBER:	60/257,495
	PRIOR FILING DATE:	2000-12-21
	PRIOR APPLICATION NUMBER:	60/258,171
	PRIOR FILING DATE:	2000-12-20
	PRIOR APPLICATION NUMBER:	60/269,940
	PRIOR FILING DATE:	2001-02-20
	PRIOR APPLICATION NUMBER:	60/274,192
	PRIOR FILING DATE:	2001-03-08
	PRIOR APPLICATION NUMBER:	60/277,826
	PRIOR FILING DATE:	2001-03-22
	PRIOR APPLICATION NUMBER:	60/279,840
	PRIOR FILING DATE:	2001-03-29
	PRIOR APPLICATION NUMBER:	60/282,981
	PRIOR FILING DATE:	2001-04-11
	PRIOR APPLICATION NUMBER:	60/283,656
	PRIOR FILING DATE:	2001-04-13
	PRIOR APPLICATION NUMBER:	60/309,247
	PRIOR FILING DATE:	2001-07-31
	PRIOR APPLICATION NUMBER:	60/311,754
	PRIOR FILING DATE:	2001-08-17
	PRIOR APPLICATION NUMBER:	60/313,331
	PRIOR FILING DATE:	2001-08-17
	NUMBER OF SEQ ID NOS:	260
	SOFTWARE:	PatentIn Ver. 2.1
	SEQ ID NO 33	
	LENGTH:	967
	TYPE:	DNA
	ORGANISM:	Homo sapiens
	US-10-032-189-33	

	Query Match	42.8%	Score 403.6;	DB 7;	Length 967;
	Best Local Similarity	65.1%;	Pred. No. 1.7e+103;		
	Matches 595;	Conservative 0;	Mismatches 319;	Indels 0;	Gaps 0;
OY	2 atgaattggtgaataagacagcatcactaacagsgattatcttgctggtttcagatoga	61			
Db	12 atggctcagaanaaaaggaggtcttccactcgagattaccctcacggygttccgcgacaag	71			
OY	62 ccttgctggaagtcccaactcccttggtgtctcttgatcttcttaacactgtagccaacctt	121			
Db	72 cctcagctggagctagctctctcttggttgttcttcttgatcttcttatatacttcaacttgctg	131			
OY	122 ggcgaatcgcagcatattcttagtgcacagcgcctggagaccaaaaattcataccccatgat	181			
Db	132 gggaacaaaacatcatgtatattatctactctggaccaccacatcttcaaatctcatgtat	191			
OY	182 tttttcttcaacactcatcactcctctggtatcttggttacacacatgtagctccacaa	241			
Db	192 ttttctcttccaactaagcttlttgatctgtgttacacaaocggatgtgtccaaag	251			
OY	242 atgctagtaaatltagtcagcatcacggaagaatlaacagtlacgtgctgtgtgaccag	301			
Db	252 ctccctgtaatatctcaggggagcagacaacaatcatctcatgtggtgtgtgattcag	311			
OY	302 ctttccatatatttgcccttgggggctactgtaatatcttctccctggcgcgtatgctctt	361			
Db	312 ctgtacatctctcttagtgcttggagatcacagaatgcttctctttagggatgattgcatct	371			
OY	362 gatgtgttgtagctattcttgctgcgcctccatctacatcagttatcatcacacaagagctc	421			
Db	372 gagcgctatgcaagctgtttgacaggccctccactacaacagtagtcatgcaacctgtctg	431			
OY	422 tgctctcaagttggcagcatcactggttactcgtgtttagttaactcaggtgtgtgtct	481			

Db 432 tatgtgctatgtgcttactatcatggtgctatgtgtttgcacactccatctgcaagc 491

QY 482 accctgactctcagctgcaactctgtaaccctatgtatagatcatcttctctgtaa 541

Db 492 gtgtcatcttctgtttaacactttgtggaagaataatagaaacacttcttctgag 551

QY 542 gtccctgcatctgtcaagttaatttgtgtgtaagaacaacgaatgagcttgaacttc 601

Db 552 gtctccatctgtgcaagctgtgctgttgaactacttgaatgaatctgaactcttc 611

QY 602 ctgttcagctgagctctcatctcaatccactgaactcaactccatataatgacttc 661

Db 612 ttgttcagctgtcatatcttctctgttaccctgttgcataatcatctctccatagtcag 671

QY 662 attgtccgaagcatatgtgaagatacagctgtctgaaggtcgcacaanaagcatltggaca 721

Db 672 attgtcagggcagcatgagataaagtcagcaacagggcaggaanaagtgtttgggaca 731

QY 722 ttgtgttccatctaatgtgtgtgtctctttttatagtaagcgcgtctgttactctg 781

Db 732 ttgtgtccctccacactcaagatgtgttccctgttctacgcagcatctatgttactctc 791

QY 782 caaccaccttcgcgcacacgaagcaagaaagtgttctctctctatggaatc 841

Db 792 cagcccgcaacaactctctcaggaatcagggcagatctctctctctatcacatc 851

QY 842 attgaccccatgtctgaatcccttatataacacttagaacaagagtaagaagagc 901

Db 852 attacaccatgtatcaaccctcatatatacactgaggaacaagatgtgaagagca 911

QY 902 tttaaaagtggt 915

Db 912 cttaagaagtgct 925

RESULT 3

US-10-032-106-2

; Sequence 2, Application US/10032106

; GENERAL INFORMATION:

; APPLICANT: Peng, Zaoyuan

; APPLICANT: Yu, Zailin

; APPLICANT: Willey, Andrew

; APPLICANT: Hu, Qianjin

; TITLE OF INVENTION: METHODS FOR IDENTIFYING G-PROTEIN

; TITLE OF INVENTION: COUPLED RECEPTORS ASSOCIATED WITH DISEASES

; FILE REFERENCE: 433112000700

; CURRENT APPLICATION NUMBER: US/10/032.106

; PRIOR FILING DATE: 2002-04-08

; PRIOR APPLICATION NUMBER: US 60/258,070

; NUMBER OF SEQ ID NOS: 10

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 2

; LENGTH: 939

; TYPE: DNA

; ORGANISM: Homo sapiens

US-10-032-106-2

Query Match 42.6%; Score 402; DB 7; Length 939;

Best Local Similarity 65.0%; Pred. No. 4.9e-103;

Matches 594; Conservative 0; Mismatches 320; Indels 0; Gaps 0;

QY 2 atgaattgtgtaaatgacagcatatcacagagttatcttcgtctggttctcagatcga 61

Db 1 atgagatcagaanaaatggaagttcttccacgtgattatcccaacgggttcttcgaagg 60

QY 62 cctgtgctgaggttccacactcctgtgtctctcttgaatttcaactgtgacatctt 121

Db 61 cctagctgagcagctagctccttctgttcttcttctctatctatcttcaacttgcgt 120

QY 122 ggcacatgcacatattcctagctgctgacgcacacaaacttcaacccatgcat 181

Db 121 gggacaacaaacatcatgttatcttccacttggaaacccacacttcaactccatgtat 180

QY 182 ttcttcttcaacaaactatcatcactcttggatcttctgttatacacaactatgtacagtcacca 241

Db 181 ttcttcttccaaacctaagcttcttggatctgtgttatacacaacgcagctgttccacag 240

QY 242 atgtctgtaaatattatgcagcatcagaagaatgaatcaatgtaatctgtgtgtgtgcccag 301

Db 241 ctctctgttattcccaaggagagcagacaatcaatccctcatgtgtgtgtgtgtgtcag 300

QY 302 ctcttcatattctggccttggggctgctactgaaatcttctccgcgcagcatctcctt 361

Db 301 ctgtacatctctccagctgtggatctacagaatgcgttctcttagagagatgtattt 360

QY 362 gatgtgtttagatattctgtcgcctctccatctacactgaatctacatgcacacagaatc 421

Db 361 gaccgcatagagatgttttggagccctccacatcaacagatgacatgcacaccctgtctg 420

QY 422 tgccctcagttggagactgtacactcctgtgttactgtgttttagtaactcagttgtgtct 481

Db 421 tatgtgtgagtggtcttacttcaatctgtgtcatltgttcttgcacactccatcttgcagagc 480

QY 482 accctgactctccagctgcacactctgtgaccccatgtgtatgatacacttctctgtgaa 541

Db 481 gtgtcatcttctgtttaaactcttgtggaagaataataatgaaacacttcttctgtgag 540

QY 542 gtccctgcatctcgaagttaattctgtgtgtgtaagaacaacagaaatgaggtgaaactatc 601

Db 541 gtctctcatctgtccaaactgtgctgtgtgtgaacactactatgaatgaatctgaaactctc 600

QY 602 ctgtcagttgagctctccatctataaccctgacactcaactcttatacatatgacttt 661

Db 601 ttgtcagttgcatatcttcttctgttactcgttgcattatcaatcttctctatagtcag 660

QY 662 attgcagcagatgatgagagatacagctgtctgaaggtcgacaanaagcatltgggaca 721

Db 661 attgcagggcagatcagagagataaagttagcaacacagggcagagaanaagtttgggaca 720

QY 722 ttgtgtcccatctaatgtgtgtctcttcttattatagtaacagcgcgtctgttactctg 781

Db 721 ttgtgtcccatctcaacagatgtgttccctgttcttcaagcagacatcatatgtcttactc 780

QY 782 caaccacacttcgcgcacacgaagcaaggaagaatggttctctcttataagaaatc 841

Db 781 cagcccgcaacaactctctcagatcaggaatgaggaatgcatctctcttctataacatc 840

QY 842 attgaccccatgtctgaatcccttatataacacttgaagaacaggaagtaagaagagc 901

Db 841 attacaccatgtatcaaccctcatatatacactgaggaacacagatgtgtgaagagca 900

QY 902 tttaaaagtggt 915

Db 901 cttaagaagtgct 914

RESULT 4

US-09-965-422-9

; Sequence 9, Application US/09965422

; GENERAL INFORMATION:

; APPLICANT: Spytek, Kimberly A

; APPLICANT: Casman, Stacie

; APPLICANT: Padigaru, Muralidhara

; APPLICANT: Dickson, Kevin

; APPLICANT: Vernet, Corine

; APPLICANT: Spaderna, Steven K

; APPLICANT: Shenoy, Suresh G

; APPLICANT: Gerlach, Valerie

; APPLICANT: Ellerman, Karen

; APPLICANT: Edinger, Shlomit

; APPLICANT: MacDougall, John R

; APPLICANT: Smithson, Glenda

; APPLICANT: Li, Li

; APPLICANT: Malyankar, Urial M


```
; CURRENT APPLICATION NUMBER: PCT/US02/03635
; CURRENT FILING DATE: 2002-02-06
; PRIOR APPLICATION NUMBER: 60/267,322; 60/271,215; 60/274,551; 60/278,507;
; 60/280,597; 60/281,107; 60/282,121
; PRIOR FILING DATE: 2001-02-07; 2001-02-23; 2001-03-08; 2001-03-23;
; 2001-03-30; 2001-04-02; 2001-04-06
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE: PERL Program
; SEQ ID NO 59
; LENGTH: 1187
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No: 7474873CB1
PCT-US02-03635-59

Query Match          39.4%; Score 371.6; DB 1; Length 1187;
Best Local Similarity 62.5%; Pred. No. 1.9e-94;
Matches 381; Conservative 0; Mismatches 349; Indels 0; Gaps 0;

QY 14 aatgacagcatcacagaagttatctctgtgttccagatcgactgtgtgag 73
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 154 aatgagagcaactagacagtttccatccttttaggtttctgtatctccagttacag 213

QY 74 ttccacctcctgtgtctctctgtattcttaacatgtgacacatttgcaatctgacc 133
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 214 aaggtcttattgttcataatgatctgtcttacttaactattttgggaataacacc 273

QY 134 atattctagtgtcagcctcgagacacaaattcacaaccccatgtattttttctacc 193
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 274 atattctgttctcgtcgtggaacccaattcatalagtcgaatgtattctctcttct 333

QY 194 aactatcatcctcgatcttcttcttcaacacatgtacagttccacaagaatgtat 253
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 334 catctctcctctctgtacacgtctctcaccagaagttatttcccccgcctcgtataac 393

QY 254 ttatgcagcatcagaagaatcaatcagttacgtgtgtgtgtaagccactttcataat 313
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 394 ctgtgggaacccaatgaataatcagctcctatgtgtgtgtgtgttcaactctcc 453

QY 314 ctggccttggggctactgatatcttctctcgtgcgtacatgttcccttggttgtgta 373
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 454 catgcccctggatccactgtgctgcctcgcgtgtgtgtgtgtgtgtgtgtgtgtgt 513

QY 374 gcatattgtgcgctccatcactcaatgattacacagagacagagacccctccagtt 433
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 514 gctgtctgcgctcctcctccttaactgtcttaatgatatccatctcctgtcagtgcc 573

QY 434 gcagctgcatcctgtgttactgtgtttagtaactcagtggtgtgtgtcttactcct 493
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 574 gcatctatgtcagtcagtcagtggaatagccacacccctgtgtacacccatccact 633

QY 494 cagctgcgaactcgtgacccctatgtatagatcacttctcctgtgtgaagccctgcact 553
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 634 cagctgcctcctcgtgtggtcgtcccaagtgatcatttcaatctcgtgaggtccctgt 693

QY 554 ctcaagtattatctgtgtgtgacacagacaagaatgtaggtctaacactcctgtgtgag 613
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 694 atcaagctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 753

QY 614 ctcttcacatlaaacctcgtgacatcactcctatatacatatgctttattgttcgagca 673
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 754 ctcttccttattagtcgtcgtcattcatctcgtctcctcctgtcattgtccacgca 813

QY 674 gtattagaggtacagtcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 733
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 814 gtgttgaggtacagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcag 873

QY 734 ctaattgtgtgtctctctttttatagtagcagcgtctctgtgtacccgacacactgtg 793
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 874 ctgacagtggtacacacttttatgtggaacacatcatcttcatgtatctgtcagccagcaag 933
```

```
QY 794 cccagctcccaagcacaagaagatgtttctctctctatagtaatcatctgacccatg 853
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 934 agtagatccaggagccagggcaagttgttctctctctcctcactgtgttaacccgcat 993

QY 854 ctgaatcccttatatacttaacttaggaacaagaggttaaggaagctttaaaggttg 913
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 994 cttaacccttattttactccttgagatcaagaaggtgtgaagggcattaaagaagtt 1053

QY 914 gtgcgaagctctcttaacaaagaataa 943
    ||||| ||||| ||||| ||||| |||||
Db 1054 ctgacaaagctcctggtggaatatttta 1083
```

```
RESULT 6
US-09-965-422-7
; Sequence 7, Application US/09965422
; GENERAL INFORMATION:
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Casman, Stacie
; APPLICANT: Padigaru, Murallidhara
; APPLICANT: Dickson, Kevin
; APPLICANT: Verneet, Corine
; APPLICANT: Spaderna, Steven K
; APPLICANT: Shenoy, Suresh G
; APPLICANT: Gerlach, Valerie
; APPLICANT: Ellerman, Karen
; APPLICANT: Edinger, Shlomil
; APPLICANT: Macdougall, John R
; APPLICANT: Smithson, Glenda
; APPLICANT: Li, Li
; APPLICANT: Malpankar, Urial M
; APPLICANT: Taylor, Sarah
; APPLICANT: Gunther, Erik
; APPLICANT: Tchemnev, Velizar T
; TITLE OF INVENTION: Novel Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21401-132
; CURRENT APPLICATION NUMBER: US/09/965,422
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: 60/236,286
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: 60/236,284
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: 60/237,581
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/238,735
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/240,736
; PRIOR FILING DATE: 2000-10-16
; PRIOR APPLICATION NUMBER: 60/260,019
; PRIOR FILING DATE: 2001-01-05
; PRIOR APPLICATION NUMBER: 60/260,338
; PRIOR FILING DATE: 2001-01-08
; PRIOR APPLICATION NUMBER: 60/262,156
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 60/262,498
; PRIOR FILING DATE: 2001-01-18
; PRIOR APPLICATION NUMBER: 60/263,133
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: 60/263,691
; PRIOR FILING DATE: 2001-01-24
; PRIOR APPLICATION NUMBER: 60/266,109
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 60/271,634
; PRIOR FILING DATE: 2001-02-26
; NUMBER OF SEQ ID NOS: 127
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 1007
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-965-422-7
```

Query Match 39.2%; Score 370; DB 5; Length 1007;
Best Local Similarity 62.4%; Pred. No. 5.2e-94;
Matches 580; Conservative 0; Mismatches 350; Indels 0; Gaps 0;

QY 14 aatgacagcatcatcagaaggttatctctgctggttcttcacagctgtgag 73
DB 75 aatgagagcaacctagaggtttcatctcttgggtttctgtatctcctcaag 134
QY 74 ttccacctctgtgtctcttctgattcttctacactgtgacatcttgcaactgacc 133
DB 135 aaggttatttctgtcccatatgattctgtattactatctatttggggaataccacc 194
QY 134 attattctagctcagctgtgacacaaactcattaccccatgtaatttttcttacc 193
DB 195 atcatctgttctctgctcggaaccagcttcatatgctcgatgtattctctcttcc 254
QY 194 aactatcatctctgtgattttgtttacacatgtacagtcaccacaaagttagtaac 253
DB 255 catctctctctctgtacccgtcttccacagacagtgatttaccacgctctgttaac 314
QY 254 ttatcgacatcaggaagaatcagttatcgtgctgtgtagccagctttcataatt 313
DB 315 ctgttggaacccatgaaactatcgtcctatggtgctgttctgttccacttaactcc 374
QY 314 ctggccttggggctactagatatctctccctggcgctacgtctcttggattgta 373
DB 375 catgcccctggatccactagtggtctcccgctctgattctctgtgacccgtatggt 434
QY 374 gctattgtggcctctccacttactcagtatcatcagacagaaactctgctcattg 433
DB 435 gctgctccgctccctccctccctacacgtcttcaatcattcgtcgtgacccgtg 494
QY 434 gcaagctcagctcgtgtgtactgtttagtaactcagtggtgtgtctaccctactc 493
DB 495 gcatctatgcatgagccatgagatagcaccacccctgtgtaacgtccacccactctg 554
QY 494 cagctgcactcgtgacccctatgtatgatacattctctgtgtaagccctgacgt 553
DB 555 cagctgcctctctgtggtgacatcgcaagtgatcatctctgctgagctccctgtctc 614
QY 554 ctcaagttatctgtgtgtggaacacagcaaaagacattggaactcctgtcagtgag 613
DB 615 atcaagctcgtcgtgtggaacacagcttcaacgagcgtgagcttctgtggtagatc 674
QY 614 ctcttcatcattacaccccgacacatcctctatcatatgctttattgtctcgagca 673
DB 675 ctcttcttattagctcgtctcatcattcatctcgtctccctgtctacattgtccca 734
QY 674 gtattgagatcacgtctgtgaagtcgacaaaagacatttgagacattgttcccat 733
DB 735 gttgtgagattagtcagtcacagagacagaaacattcgtgacctgtctctccac 794
QY 734 cttaattgtgtgtctcttttttattagtaacgctgtctgttctacatcgacaacctg 793
DB 795 ctgacagtggttcacacatttcttggaacacatcattcttcgtcgtcagccagca 854
QY 794 cccagctcccaaggaacaaagatggttctctcttcatatgaaacatctgccccatg 853
DB 855 agtagatccagggagcccggaagtgttctctctctcttcaactgtgtaaccccgtag 914
QY 854 ctgaatccctatataatacacttaggaacaggaaggttaagaaagcctttaaaggtg 913
DB 915 cttaacaccttattatatacctttagagatcaagaaggtgaaagggatctaaagaagtc 974
QY 914 gttgcaagatcttcttcaatcaagaataa 943
DB 975 ctgcaagaagctctggaagtaaatatttta 1004

RESULT 7
US-09-965-422-3
; Sequence 3, Application US/09965422

GENERAL INFORMATION:
APPLICANT: Spytek, Kimberly A
APPLICANT: Casman, Stacie
APPLICANT: Padigar, Muralidhara
APPLICANT: Dickson, Kevin
APPLICANT: Vernet, Corine
APPLICANT: Spaderna, Steven K
APPLICANT: Shenoy, Suresh G
APPLICANT: Gerlach, Valerie
APPLICANT: Ellerman, Karen
APPLICANT: Edinger, Shlomit
APPLICANT: MacDougall, John R
APPLICANT: Smitson, Glenda
APPLICANT: Li, Li
APPLICANT: Malvanekar, Urial M
APPLICANT: Taylor, Sarah
APPLICANT: Gunther, Erik
APPLICANT: Tchernev, Velizar T
TITLE OF INVENTION: Novel Proteins and Nucleic Acids Encoding Same
FILE REFERENCE: 21401-132
CURRENT APPLICATION NUMBER: US/09/965,422
CURRENT FILING DATE: 2001-09-27
PRIOR APPLICATION NUMBER: 60/236,286
PRIOR FILING DATE: 2000-09-28
PRIOR APPLICATION NUMBER: 60/236,284
PRIOR FILING DATE: 2000-09-28
PRIOR APPLICATION NUMBER: 60/237,581
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/238,735
PRIOR FILING DATE: 2000-10-06
PRIOR APPLICATION NUMBER: 60/240,736
PRIOR FILING DATE: 2000-10-16
PRIOR APPLICATION NUMBER: 60/260,019
PRIOR FILING DATE: 2001-01-05
PRIOR APPLICATION NUMBER: 60/260,338
PRIOR FILING DATE: 2001-01-08
PRIOR APPLICATION NUMBER: 60/262,156
PRIOR FILING DATE: 2001-01-17
PRIOR APPLICATION NUMBER: 60/262,498
PRIOR FILING DATE: 2001-01-18
PRIOR APPLICATION NUMBER: 60/263,133
PRIOR FILING DATE: 2001-01-19
PRIOR APPLICATION NUMBER: 60/263,691
PRIOR FILING DATE: 2001-01-24
PRIOR APPLICATION NUMBER: 60/266,109
PRIOR FILING DATE: 2001-02-02
PRIOR APPLICATION NUMBER: 60/271,634
PRIOR FILING DATE: 2001-02-26
NUMBER OF SEQ ID NOS: 127
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 3
LENGTH: 1013
TYPE: DNA
ORGANISM: Homo sapiens
US-09-965-422-3

Query Match 39.2%; Score 370; DB 5; Length 1013;
Best Local Similarity 62.4%; Pred. No. 5.2e-94;
Matches 580; Conservative 0; Mismatches 350; Indels 0; Gaps 0;

QY 14 aatgacagcatcatcagaaggttatctctgctggttcttcacagctgtgag 73
DB 81 aatgagagcaacctagaggtttcatctcttgggtttctgtattctcctcagttacag 140
QY 74 ttccacctctgtgtctcttctgattcttctacactgtgacatcttgcaactgacc 133
DB 141 aaggttatttctgtcccatatgattctgtattactatctatttggggaataccacc 200
QY 134 attattctagctcagctgtgacacaaactcattacaccccatgtaatttttcttacc 193
DB 201 atcatctgttctctgctcggaaccacagcccatcatatgctcgatgtattctctcttcc 260

```

QY 194 aatcatcactcctcgatcttctgttaacacacatgtaacgtccacaaatgctaataat 253
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 261 catctcctcctcctgcttcacccgtcttcacccagctgttattcccgagctgtgtaaac 320
QY 254 ttatgcagcatcaagaagaatcagctatcgtcgtgtgtagccagctttcatattt 313
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 321 ctgtgggaacccacgaagaataatcgcctatggtggtctgttctgtaaccttaaacccc 380
QY 314 ctggccttgggggctactgaataatctctcctggccgctgactgtccttgatgttgta 373
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 381 catgcctggatccactgagtgctgctcccgctcctgagctgctgtgacgctagtgtg 440
QY 374 gctattgtgcgcctccctcacttaactcaatgataatgacacagagactgcctcagttg 433
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 441 gctgtccgcctcctcctccctacacgtcttaactgaatcatalccacttgcagtgcctg 500
QY 434 gcaagctgcatccctggttactgagtttagtaactcagtggtgtgtcttaccctgaacttc 493
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 501 gcatctatgcatgtgtcctcagtgtaataagccacccctgtgtgtaagcttccacccctc 560
QY 494 caagctgcacactgtgtgaacccctatgtatagatacacttctctgttgaagctccctgac 553
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 561 cagctgcctctctgtgggacatgcagtgtaacatcttcatctcgtgagctccctgtgctc 620
QY 554 ctcaagttactctgtgttgagacaacagcaaatgagctgtgacacttccctgtcagtgag 613
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 621 atcaagctgtgtgtgtgggacacacgtttaaagagctgtgagcttctgtgtcagtaatc 680
QY 614 ctctccatctaaacacccctgacacatcccttcatcattatgtcttattgtctgcagaca 673
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 681 ctttctcttatagtcgtgtctcacttcatcctcgtctcctcctgtctaatgtgccaaagca 740
QY 674 gtattgagabacagctcgtcgtgaagctgcacaaaagcaatlttgagacatgtgttcccat 733
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 741 ggtgtgagagtaagtaagcagctacccggagagacaagaagcattcggagacctgtctcc 800
QY 734 ctattgtgtgtctcttttttataagtaacacgctctcctgtgtgtaacctgcaacccacttc 793
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 801 ctgacagtggtgcacacatcttcttctgaacacatcttcatcgtatctgtgacgccaag 860
QY 794 cccagctcccaaggacgaagaagatggttctctctctctctatgtgaatcatgacccacatg 853
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 861 agtagatccacaggacagggacgaagtgttctctctctctctacactgtgtgtacccgcag 920
QY 854 ctgaatcccttataataacttagaacaagaggtaaaggaagcctttaaaggtgtg 913
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 921 cttaacctcttattataacttgaagatcagagagtggaaggggacatlaagaaggtt 980
QY 914 gtgcaagagctctcttaatacaagaataa 943
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 981 ctgacaaagctcgtggagtaataatttta 1010

```

```

: APPLICANT: Tchernev, Velizar T
: TITLE OF INVENTION: Novel Proteins and Nucleic Acids Encoding Same
: FILE REFERENCE: 21401-132
: CURRENT APPLICATION NUMBER: US/09/965,422
: CURRENT FILING DATE: 2001-09-27
: PRIOR APPLICATION NUMBER: 60/236,286
: PRIOR FILING DATE: 2000-09-28
: PRIOR APPLICATION NUMBER: 60/236,284
: PRIOR FILING DATE: 2000-09-28
: PRIOR APPLICATION NUMBER: 60/237,581
: PRIOR FILING DATE: 2000-10-03
: PRIOR APPLICATION NUMBER: 60/238,735
: PRIOR FILING DATE: 2000-10-06
: PRIOR APPLICATION NUMBER: 60/240,736
: PRIOR FILING DATE: 2000-10-16
: PRIOR APPLICATION NUMBER: 60/260,019
: PRIOR FILING DATE: 2001-01-05
: PRIOR APPLICATION NUMBER: 60/260,338
: PRIOR FILING DATE: 2001-01-08
: PRIOR APPLICATION NUMBER: 60/262,156
: PRIOR FILING DATE: 2001-01-17
: PRIOR APPLICATION NUMBER: 60/262,498
: PRIOR FILING DATE: 2001-01-18
: PRIOR APPLICATION NUMBER: 60/263,133
: PRIOR FILING DATE: 2001-01-19
: PRIOR APPLICATION NUMBER: 60/263,691
: PRIOR FILING DATE: 2001-01-24
: PRIOR APPLICATION NUMBER: 60/266,109
: PRIOR FILING DATE: 2001-02-02
: PRIOR APPLICATION NUMBER: 60/271,634
: NUMBER OF SEQ ID NOS: 127
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 5
: LENGTH: 1014
: TYPE: DNA
: ORGANISM: Homo sapiens
: US-09-965-422-5

Query Match          39.2%; Score 370; DB 5; Length 1014;
Best Local Similarity 62.4%; Pred. No. 5,2e-94;
Matches 580; Conservative 0; Mismatches 350; Indels 0; Gaps 0;

QY 14 aatgacgacatcaacaagaattatctcgtggttctcagatcgacacttgctgag 73
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 82 aatgagacaacttagcaggttcatcctttaggttcttgatctatccatatttgggaataccac 141
QY 74 ttccacacttggtgtcttcttgattcttaactgtgacactgtgacacatttggaacttgacc 133
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 142 aaggttcatattgtgtcatalatgtattgtattactaactatttgggaataccac 201
QY 134 attattcagtgacgcgcctgacacacaaacttcataccocatgtatttttctacc 193
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 202 atcatctcgttctcgtctgtgaacccaagcttcatatgcccagatgatcttcttcttct 261
QY 194 aatcatcactcgtgacatcttgttgaacacaaatgtacagttccacaaatgctagtaaat 253
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 262 catctcctcctcgttaccgctgttccacagcaggttatttccacagctcctggttaaac 321
QY 254 ttatgcagcatcagaagaatcagttatcgtatcgtgtgtgtagccagctttcatattt 313
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 322 ctgtgggaacccatgaagaataatcgcctatggtggtctgttctgtaaccttaaacccc 381
QY 314 ctggccttgggggctactgaataatctctcctggccgctgactgtccttgatgttgta 373
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 382 catgcctgggataccagagtgctcctcccgctcctgagctgctgtgacgcagtg 441
QY 374 gctattgtgcgcctcctcacttaactcagttatcatgacacagagactgcctcagttg 433
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 442 gctgtctgcgtctcctcacttaacactgtcttaatgacatacatcctctgtatggccttg 501
QY 434 gcaagctgcatccctggttactgagtttagtaactcagtggtgtgtcttaccctgaacttc 493

```

Db 502 gcaactatgacatggtcctagctggaatgacccaccaccggtgacagcttcacccctaaccttg 561
Qy 494 cagctgcacacctctgacccctatgtaatacattctctctggaagctccctgac 553
Db 562 cagctgcacacctctggaacacgacgacgaatgatacttctcagagctccctgctc 621
Qy 554 ctcaagttatctggtgtgagacacagcaaatgaggtcgaactatctctgtcagtgag 613
Db 622 atcaagctggtgtgtgagacacagctttaaaggtcgtgacttttgtgtcagtgatc 681
Qy 614 ctcttcacatcaataacccctgacacatcctctatatacattgctttatgtccgagca 673
Db 682 ctcttcacatcaataacccctgacacatcctctatatacattgctttatgtccgagca 741
Qy 674 gtattgagatcacagctctgctgaaggtcgaacaaagacattggtgagacatgtgtccat 733
Db 742 gtgttgagatcaataagctcagacacagagagacagaagattcggagactgtctctccac 801
Qy 734 ctactgtgtgtctctctttttatagtaacagcgtctctgtgtacccgacacacccctcg 793
Db 802 ctgacagtggtcaccacattttatggaacacacatcctcatctgtatctgacgacgacag 861
Qy 794 cccagctcccaagagcaagaagaatgttctctctctctctatggaatcatgtcacccatg 853
Db 862 agtagatccagagacagcaggaagttgttctctctctctctctacactgtgtacccgacatg 921
Qy 854 ctgaaccccttatatacacttaggacacaggaaggaaggttaaaaggtg 913
Db 922 cttaaccccttatatacacttaggacacaggaaggttaaaaggtgacataaagaagtt 981
Qy 914 gtgcagaagctctcttaatacaagaataa 943
Db 982 ctgaagaagctctggaagtaataatttta 1011

RESULT 9
US-09-546-986A-7
; Sequence 7, Application US/09546986A
; GENERAL INFORMATION:
; APPLICANT: Powers, Scott
; APPLICANT: Yang, Jianxin
; APPLICANT: Cutler, Gene
; APPLICANT: Tularik Inc.
; TITLE OF INVENTION: Novel G-Protein Coupled Receptors
; FILE REFERENCE: 018781-004720US
; CURRENT APPLICATION NUMBER: US/09/546, 986A
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 09/524, 730
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 7
; LENGTH: 1065
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (26)..(1030)
; OTHER INFORMATION: human breast cancer amplified G-protein coupled
; OTHER INFORMATION: receptor 4 (BCA-GPCR-4)
US-09-546-986A-7

Query Match 39.2%; Score 370; DB 5; Length 1065;
Best Local Similarity 62.4%; Pred. No. 5.3e-94;
Matches 580; Conservative 0; Mismatches 350; Indels 0; Gaps 0;

Qy 14 aatgacagcatcatacaggaattatctctggtggttcacagatcgacctgtgag 73
Db 98 aatgagagcaaccagaggtttcatctctttaggtttctgttatgtcagttacag 157
Qy 74 ttccactcctgtgtctctctgtattcttaacatgtgacacatcttgcaatctgacc 133

Db 158 aaggtctattctgtgtcattatattctgtatttacttaactatttgggaataaccacc 217
Qy 134 attattcagtgtaacgcttgacacacaaacttaacttaacccatattttttcttacc 193
Db 218 atcaattcgtttctctgtcctggaacacagcttcaatagccgaagtattcttctctct 277
Qy 194 aatcatcactcctggaatcttctgttacacacacatgtacagtcacacaaatgtagtaac 253
Db 278 catctctctctctgtacacgctgtcttaacagagtgattatctccacagctctctgtaac 337
Qy 254 ttatgacacatcaaggaagatacagttatctgtgtgtgtgtgagccagctttcataatt 313
Db 338 ctgttggaacacatgaacacatcgcctatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 397
Qy 314 ctgacgttggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtg 373
Db 398 catgacctggtatccaactgagtggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 457
Qy 374 gctattgtggtcctctccatctacatgtaacatgacacacacacacacacacacacacac 433
Db 458 gtgtgtggtcgtctctccatctacatgtaacatgacacacacacacacacacacacacac 517
Qy 434 gcagctgacatctgt 493
Db 518 gcatctatgacatgt 577
Qy 494 cagctgcacacctctgacccctatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 553
Db 578 cagctgcacacctctgacccctatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 637
Qy 554 ctcaagttatctgt 613
Db 638 atcaagctggt 697
Qy 614 ctcttcacatcaataccctgac 673
Db 698 ctcttcacatcaataccctgac 757
Qy 674 gtattgagatcacagctctgac 733
Db 758 gtgttgagatcaataccctgac 817
Qy 734 ctactgtgtgtctctctttttatagtaacagcgtctctgtgtgtgtgtgtgtgtgtgtgt 793
Db 818 ctgacagtggtcaccacatctttatggaacacacacacacacacacacacacacacacac 877
Qy 794 cccagctcccaagagcaagaagaatgttctctctctctctctctctctctctctctctct 853
Db 878 agtagatccagagcagcaggaagttgttctctctctctctctctctctctctctctctct 937
Qy 854 ctgaaccccttatatacacttaggacacaggaaggttaaaaggtttaaaggtt 913
Db 938 cttaaccccttatatacacttaggacacaggaaggtttaaaggtttaaaggtt 997
Qy 914 gtgcagaagctctcttaatacaagaataa 943
Db 998 ctgaagaagctctggaagtaataatttta 1027

RESULT 10
US-09-965-422-11
; Sequence 11, Application US/09965422
; GENERAL INFORMATION:
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Gasman, Stacie
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Dickson, Kevin
; APPLICANT: Vernet, Corine
; APPLICANT: Spaderna, Steven K
; APPLICANT: Shenoy, Suresh G
; APPLICANT: Gerlach, Valerie
; APPLICANT: Ellerman, Karen

```

; APPLICANT: Edinger, Shlomit
; APPLICANT: MacDougall, John R
; APPLICANT: Smithson, Glenda
; APPLICANT: Li, Li
; APPLICANT: Malvaner, Uriel M
; APPLICANT: Taylor, Sarah
; APPLICANT: Gunther, Erik
; APPLICANT: Tcheney, Velizar T
; TITLE OF INVENTION: Novel Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21401-132
; CURRENT APPLICATION NUMBER: US/09/965,422
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: 60/236,286
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: 60/236,284
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: 60/237,581
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/238,735
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/240,736
; PRIOR FILING DATE: 2000-10-16
; PRIOR APPLICATION NUMBER: 60/260,019
; PRIOR FILING DATE: 2001-01-05
; PRIOR APPLICATION NUMBER: 60/260,338
; PRIOR FILING DATE: 2001-01-08
; PRIOR APPLICATION NUMBER: 60/262,156
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 60/262,498
; PRIOR FILING DATE: 2001-01-18
; PRIOR APPLICATION NUMBER: 60/263,133
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: 60/263,691
; PRIOR FILING DATE: 2001-01-24
; PRIOR APPLICATION NUMBER: 60/266,109
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 60/271,634
; PRIOR FILING DATE: 2001-02-26
; NUMBER OF SEQ ID NOS: 127
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO: 11
; LENGTH: 1005
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-965-422-11

Query Match      37.9%; Score 357; DB 5; Length 1005;
Best Local Similarity 61.4%; Pred. No. 2.4e-90;
Matches 573; Conservative 0; Mismatches 360; Indels 0; Gaps 0;

```

```

Db 355 cttatattctctgcaactggcctcccaatgatactctctgctgacatgacctg 414
Qy 362 gattggtttgtagactatttgcgcctctccatctacatgatactgacacagacac 421
Db 415 gatcggtaacttgctgtctgcaaacccctccactatgtagtaactgaacacagctt 474
Qy 422 tgcctcagttgcaactgcaactgctggttactggttttagtaactgagtggtgtct 481
Db 475 tgcacaagctggaactatctctctgctcagtggtttggtgattccctacacagca 534
Qy 482 accctgactcagctgcaactctgtgaccctatgtagatgatacttctctgtgaa 541
Db 535 acccttaactgcaactgctctctctgtgcaacacatagctggaacatttatttgcga 554
Qy 542 gtccctgcaactgcaactgatacttctgtttagacaacagcaaatgagctgcaactac 601
Db 595 gtacacagctctcgaagtgtgtgtgtgtagacacacagctgcaatgagtgctttt 654
Qy 602 ctgtcaagtgagctctccatcaataccctgacacatccttatacatgctttt 661
Db 655 gtgtgtagtgctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 714
Qy 662 atgtccgaagcagattggaagatacagctgtgtggaagtgacaaagaattgggaca 721
Db 715 ataaccaagctgtgtgtgagatcaaatcagtagaggaagacataaagctccagacc 774
Qy 722 tgtgtgtcccatctaatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 781
Db 775 tgtctctccacacttaacagtgagatatactctatgacacataactctagtgta 834
Qy 782 caaccacactgcccagctcccaaggaaggaagatgtttctctctctatgaaac 841
Db 835 caactcagtgacagctatgcccaggaaggaatttctccctctctacacatg 894
Qy 842 atgcaacccatgctgtaaccccttatatacatcactgaggaaggaaggaagc 901
Db 895 gtaccccccaacttaactcctatcactaacttaaggaacagataagaagagct 954
Qy 902 tttaaaagttgtgtgcaagagcttcttatac 934
Db 955 ctgaggaactctctctcggaataattgtgttc 987

RESULT 11
PCT-US02-03635-80
; Sequence 80, Application PC/TUS0203635
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: TRIBOULEY, Catherine M.
; APPLICANT: NGUYEN, Daniel B.
; APPLICANT: THORNTON, Michael
; APPLICANT: YAO, Monique G.
; APPLICANT: KALLICK, Deborah A.
; APPLICANT: GANDHI, Ameena R.
; APPLICANT: WALIA, Nalinder K.
; APPLICANT: ARVIZU, Chandra
; APPLICANT: ELLIOTT, Vicki S.
; APPLICANT: HAFALIA, April J.A.
; APPLICANT: RAMKUMAR, Jayalaxmi
; APPLICANT: JIN, Pei
; APPLICANT: TANG, Y. Tom
; APPLICANT: YUE, Henry
; APPLICANT: REDDY, Roopa
; APPLICANT: BURFORD, Neil
; APPLICANT: LU, Dung Anna M.
; APPLICANT: GRAUL, Richard C.
; APPLICANT: KAHN, Farrah A.
; APPLICANT: WALSH, Roderick T.
; APPLICANT: ISON, Craig H.
; APPLICANT: RICHARDSON, Thomas W.
; APPLICANT: GRIFFIN, Jennifer A.
; APPLICANT: WARREN, Bridget A.

```

APPLICANT: YANG, Junming
APPLICANT: LEE, Ernestine A.
APPLICANT: HARLAND, Lee
TITLE OF INVENTION: G-PROTEIN COUPLED RECEPTORS
FILE REFERENCE: PI-0357 PCT
CURRENT APPLICATION NUMBER: PCT/US02/03635
CURRENT FILING DATE: 2002-02-06
PRIOR APPLICATION NUMBER: 60/267,322; 60/271,215; 60/274,551; 60/278,507;
60/280,597; 60/281,107; 60/282,121
PRIOR FILING DATE: 2001-02-07; 2001-02-23; 2001-03-08; 2001-03-23;
2001-03-30; 2001-04-02; 2001-04-06
NUMBER OF SEQ ID NOS: 96
SOFTWARE: PERL Program
SEQ ID NO 80
LENGTH: 1124
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
OTHER INFORMATION: Incyte ID No: 7477359CB1
PCT-US02-03635-80

Query Match 37.9%; Score 357; DB 1; Length 1124;

Best Local Similarity 61.4%; Pred. No. 2,5e-90;

Matches 573; Conservative 0; Mismatches 360; Indels 0; Gaps 0;

QY 2 ataattggtgtaataacagacatcacatagaggttatctctgtgtgttcagatcga 61
DB 101 atgggattgggcaataagagttcccaatgattccatctcttaagcttcagacc 160
QY 62 ccttggtcgtgagttccactccttggtctctcttgaattcttacaactgacatctt 121
DB 161 cctgcgtcgtgaggtctctctcttgaattcttgaattcttcaactcctgacatctg 220
QY 122 ggaactcgtacattatctgaagtcacgcttgacacccaattacatcccatgata 161
DB 221 ggaactcgtacattatctgaagtcacgcttgacacccaattacatcccatgata 280
QY 182 tttttctccacatcatcatcctcgtgattgtttacacacatgatacgtccca 241
DB 281 tttttctccacatcatcctcgtgattgtttacacacatgatacgtcccatgata 340
QY 242 atgctgtaataattatgacagcatcaggaagtaacatgatacgtgctgtgagccag 301
DB 341 accctagtaactcgtacagacccaagacagcatcacttaacgtgtgtgtgtgacaa 400
QY 302 ctcttcatattctggtcctgggtggtgactgaatctctctctggtcgtgacgtcctt 361
DB 401 ctcttcatattctggtcctgggtggtgactgaatctctctctggtcgtgacgtcctt 460
QY 362 gattgtttgttagctattgttcggtcctccatcttactcaagtaacatgacacagagac 421
DB 461 gattgtttgttagctattgttcggtcctccatcttactcaagtaacatgacacagag 520
QY 422 tgcctcagttgtgacatcgtacccgtgtgattggtttgaactcagatggtgtgtct 461
DB 521 tgcctcagttgtgacatcgtacccgtgtgattggtttgaactcagatggtgtgtct 580
QY 482 acccttaccctcagctcgtacacatctgtgacccatgattgatacattctctctga 541
DB 581 acccttaccctcagctcgtacacatctgtgacccatgattgatacattctctctga 640
QY 542 gtccctcagctcgtacacatctgtgatacaccatgacacagaaatgaggtgacatc 601
DB 641 gtccctcagctcgtacacatctgtgatacaccatgacacagaaatgaggtgacatc 700
QY 602 ctgtcagtgaggtccttccatcacaacccctgacacatccttatacatgctttt 661
DB 701 gtgtcagtgaggtccttccatcacaacccctgacacatccttatacatgctttt 760
QY 662 atgttcgagcagttatgtgagatcagctgtcgaaggtcgcacaaaagcatgttgaca 721

DB 761 ataactaagctgtgtcgtgagatcaaatcagtagaggtcgaagcacaagccttcagacc 820
QY 722 tgtgttcccatcattatgtgtgtctctcttttatagtaacgacctgtgtgacctg 761
DB 821 tgtctctcccatcattatgtgtgtctctcttttatagtaacgacctgtgtgacctg 880
QY 782 caaccacactcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 841
DB 881 caaccacactcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 940
QY 842 atgcacacatcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 901
DB 941 gtgaccccatcattatcctcattatcattatcattatcattatcattatcattatcatt 1000
QY 902 tttaaaaggtgtgtggaaggtccttcttaac 934
DB 1001 ctgaggaactctctcgtggaaggtccttcttaac 1033

RESULT 12

US-09-546-986A-1

Sequence 1, Application US/09546986A

GENERAL INFORMATION:

APPLICANT: Powers, Scott

APPLICANT: Yang, Jianxin

APPLICANT: Cutler, Gene

APPLICANT: Tularik Inc.

TITLE OF INVENTION: Novel G-Protein Coupled Receptors

FILE REFERENCE: 018781-004720US

CURRENT APPLICATION NUMBER: US/09/546,986A

CURRENT FILING DATE: 2002-04-30

PRIOR APPLICATION NUMBER: US 09/524,730

PRIOR FILING DATE: 2000-03-14

NUMBER OF SEQ ID NOS: 16

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 1

LENGTH: 1035

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: CDS

LOCATION: (42)..(974)

OTHER INFORMATION: human breast cancer amplified G-protein coupled

OTHER INFORMATION: receptor 1 (BCA-GPCR-1)

US-09-546-986A-1

Query Match 36.6%; Score 345; DB 5; Length 1035;

Best Local Similarity 61.5%; Pred. No. 5.8e-87;

Matches 352; Conservative 0; Mismatches 345; Indels 0; Gaps 0;

QY 30 aggaattatctcgtcgtgttctcagatcagacttgctggtgttccactcctgtg 89
DB 61 aagccttaccctctcgtggtgtgtgtcgaagcgtggtcgtggaacccctcttgg 120
QY 90 tctcttgaattcttcaactgtgacatcttggcaatcgtacatattatctgtac 149
DB 121 tctcttgaattcttcaactgtgacatcttggcaatcgtacatattatctgtac 180
QY 150 gcttgacacacacatcattaccccatgattttcttcaacatctcaactcctg 209
DB 181 ggttgatcctcaactcacaagcccatgatacttctcctcagtaacacctgtccttctg 240
QY 210 atcttgttacacacatgatacagtcacacaaatgctagtaaatattatgacatcaga 269
DB 241 acctcgttac 300
QY 270 aagtaatcagttatcgt 329
DB 301 agaccctcagctatgaggt 360
QY 330 ctgaatatctctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 389

Db 361 cggagtgatcgtcctggccgcatgacctgacccgtacgtgcccagctgcaagccc 420
Qy 390 tccatactacagttatcatgacacagactcgtccacgttggaagctgcactcgg 449
Db 421 tgcatactccgtctcatcaccgctgctctcgtcagcagcctgctcgtcggc 480
Qy 450 ttacgtgtttagtaactcagtggtgtgtctacccctgacatccacgtcctcgt 509
Db 481 tcagtggtctcgcaactcctcgtcgaagtggtcccgagtgcaattgcacatccg 540
Qy 510 acccctatgtatagatcactctcctgtgaagtcctcgcagcagcctcaattctgt 569
Db 541 ggcgcagagtgctgcaacaactttctgtgagtgccgcgcagtgatcaagctcgt 600
Qy 570 ttgagcaacagcaaatgagtggaactctcctgtcagtgagctctccactaac 629
Db 601 ctgacacgcgtatgaatgacacacactgctgtcgtgtgacctctcgtgtgtgc 660
Qy 630 ccttgacactcactatcatatgacttttatgttcgagagagatgtgagatacag 689
Db 661 ccttgctcctcactcctctcctcctatggtcttatgtccgcgagtgatcagatcag 720
Qy 690 ctgctgaagtgcaacaagaacatttgagacatgtgtccactcaattgtgtctc 749
Db 721 ccccaagagagacacaaagccttgggagcgtgtccctccacgtatgatcgtccc 780
Qy 750 ttttatagtacagcgtctcgtgtacctgcaaccacctgcgccagctcaagacc 809
Db 781 tctctacactcactcgtatcatatcatgtatctgagccccctccagttactccaag 840
Qy 810 aaggaagatggttctctctctatgaaatcttgacacacagctgaaatcccttat 869
Db 841 agggcaaatctatctctctctatccataacacccacactctcaatccctcct 900
Qy 870 atcacttgaagcaaggaagtaaggaagctttaaagaattgtgttgcagaagct 926
Db 901 acacccctgagaataagatagaagggtcgtctgaggaactctgtgccagatct 957

RESULT 13
PCT-US02-03635-96
; Sequence 96, Application PC/TUS0203635
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: BUGHN, Mariah R.
; APPLICANT: TRIBOULEY, Catherine M.
; APPLICANT: NGUYEN, Daniel B.
; APPLICANT: THORNTON, Michael
; APPLICANT: YAO, Monique G.
; APPLICANT: KALLICK, Deborah A.
; APPLICANT: GANDHI, Ameena R.
; APPLICANT: WALIA, Nandier K.
; APPLICANT: ARVIZU, Chandra
; APPLICANT: ELLIOTT, Vicki S.
; APPLICANT: HAFALIA, April J.A.
; APPLICANT: RAMKUMAR, Jayalakmi
; APPLICANT: JIN, Pei
; APPLICANT: TANG, Y. Tom
; APPLICANT: YUE, Henry
; APPLICANT: REDDY, Roopa
; APPLICANT: BURFORD, Neil
; APPLICANT: LU, Dzung Alina M.
; APPLICANT: GRAUL, Richard C.
; APPLICANT: KAHN, Farrah A.
; APPLICANT: WALSH, Roderick T.
; APPLICANT: ISON, Craig H.
; APPLICANT: RICHARDSON, Thomas W.
; APPLICANT: GRIFFIN, Jennifer A.
; APPLICANT: WARREN, Bridget A.
; APPLICANT: YANG, Junning
; APPLICANT: LEE, Ernestine A.
; APPLICANT: HARLAND, Lee
; TITLE OF INVENTION: G-PROTEIN COUPLED RECEPTORS

; FILE REFERENCE: PI-0357 PCT
; CURRENT APPLICATION NUMBER: PCT/US02/03635
; CURRENT FILING DATE: 2002-02-06
; PRIOR APPLICATION NUMBER: 60/267,322; 60/271,215; 60/274,551; 60/278,507;
; 60/280,597; 60/281,107; 60/282,121
; PRIOR FILING DATE: 2001-02-07; 2001-02-23; 2001-03-08; 2001-03-23;
; 2001-03-30; 2001-04-02; 2001-04-06
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE: PERL Program
; SEQ ID NO 96
; LENGTH: 2197
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: 55019501CB1
PCT-US02-03635-96

Query Match 36.4%; Score 343.2; DB 1; Length 2197;
Best Local Similarity 61.3%; Pred. No. 2,3e-86;
Matches 552; Conservative 0; Mismatches 348; Indels 0; Gaps 0;

Qy 22 catcatagagagattatctcgtggtttctcagatcgacccctggtgagttccact 81
Db 288 caaccacaagagacttaccctccggtgcatctccgaagagcagtgctgagctccag 347
Qy 82 cctgtgtctctctgttatcttctacactgtgacatctcttgacaatctgacatct 141
Db 348 cttagagctcctcctcgtgtgtctacatcttctgcatgtcgtggaacatctctac 407
Qy 142 agtgcacgcctgagacacaacttcataccacatgtatcttttcttacaatctac 201
Db 408 ggtatcccaactggtatcctcagctccacagcccatgtatcatatctcgcgcact 467
Qy 202 actcctgactctgttgaacacacatgtacagctccacaagtgtagtaattatgag 261
Db 468 ctccctgactcctcgtacacacacacacacacacacacacacacacacacacac 527
Qy 262 catcaagaaagtaactgactcgtgtgtgtgtagccagctttcatattctgacct 321
Db 528 ctccaggaagacatcatcagtgatcgtgtgacagtgacaaagccatttccactg 587
Qy 322 gggggtactgataatctctcgtgcgtgcacatgtccttgattgtgttagctatt 381
Db 588 gggctgacagagtgctgtcttctggtgcacatgtcctgacagctacgtgcacat 647
Qy 382 tcggcctcctcatcactcagttatcatcagcaccagaagctcgtccctcagttgag 441
Db 648 tgaagcaactcgcgtatcatcatcagcaccgcccactcgtccagagcgtcgtcat 707
Qy 442 atcctggttactgtgtttagtaactcagtggtgttcttcaactcgtcctcagctgc 501
Db 708 ggcctggtcagcgtcgttggaactcctcctgttcaagtgatcctgacagtgaaat 767
Qy 502 actcgtgacccctatgtatagatcaacttctcgtggaagtcctcgtcactgct 561
Db 768 ttctgtgcccaggtgtcgtgaacaactcttctcgtgagtgccagcagtgatcag 827
Qy 562 atctgtgttgagacacagcaaatgagcgtgaactatctcctgtcagtgagcttcca 621
Db 828 gtccctgctgatactcgcggaatgagccacccctgctgtgtcgtgtgcttcttct 887
Qy 622 tctaatccctcgtacactcatcttatacatatgcttttatgtccgagcagatgtag 681
Db 888 gctgtcctcctcgtccctcctcctcctcctcctcctcctcctcctcctcctcct 947
Qy 682 gatacagctcgtcgaagtgacaaaagcaatctgggacatgtgttccatcaattgt 741
Db 948 gatcagctcctccagggagcagcagcgtcctgagactgttcttccacactgtgt 1007
Qy 742 ggtgtccttttatagtacagcgtcgtgttactcgaacacactcgcacagctc 801

Db 1008 ggtccctctcttaccctgcccacattacatgtaacctgcaagccgacatccagttac 1067
Qy 802 caaggaccagaagaatggtttctctctctctatgaaatattgacaccatctcgaaatcc 861
1068 aacggagacaggaagtaactctccctctctctctctctctctctctctctctctctcc 1127
Qy 862 ccttataatatacattaggaacaagaggtaaaggaagctttaaaaggttggttgcaag 921
Db 1128 ttcatctacacacttgaggataagagcgtgaaggagctctccgaagactcctgcaag 1187

RESULT 14

US-09-546-986A-5
; Sequence 5, Application US/09546986A
; GENERAL INFORMATION:
; APPLICANT: Powers, Scott
; APPLICANT: Yang, Jianxin
; APPLICANT: Cutler, Gene
; APPLICANT: Tularik Inc.
; TITLE OF INVENTION: Novel G-Protein Coupled Receptors
; FILE REFERENCE: 018781-004720US
; CURRENT APPLICATION NUMBER: US/09/546,986A
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 09/524,730
; PRIOR FILING DATE: 2000-03-14
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 5
; LENGTH: 1351
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (86)..(1108)
; OTHER INFORMATION: human breast cancer amplified G-protein coupled
; OTHER INFORMATION: receptor 3 (BCA-GPCR-3)
US-09-546-986A-5

Query Match 33.9%; Score 319.6; DB 5; Length 1351;
Best Local Similarity 59.8%; Pred. No. 9e-80;
Matches 535; Conservative 0; Mismatches 359; Indels 0; Gaps 0;

Qy 2 atgaattggtaaatgacgacatcaatagaggttattctgtggtttctcagatcga 61
Db 149 atgaaatagacacatgtaagtctccagaagctttgtctcctccggctctccgcaga 208
Qy 62 ccttgctggaagttccaccccttggtctcttgattcttaacactgtaacatcttc 121
Db 209 cctcactagaacatgctcctctcaatagttgctctgagtttcaacatgtaacatcttc 268
Qy 122 ggaactctgacattatctagtgacgctgacacacaaacttaccatcccatgtat 181
Db 269 ggaactgacatcatctcgtctccacataagatgtaagactccacacactatgtac 328
Qy 182 tttttcttccaatatatactcctgagatcttgtttacacacatgtacagttccacaa 241
Db 329 tttcttcttccaaacctctcctctctgagatagcttccacacagacatgttccacag 388
Qy 242 atgctagttaattatgacgacatcagaagaagtatcagttatggtgctgtgacccag 301
Db 389 ctctctgcttaacctctgaggacacagaaacataagcatgtaagggtgtgtgtcag 448
Qy 302 cttctcatattcttgacctgaggagctacgaatatctctcctgagcgtcagttcctt 361
Db 449 ttctataatcccatgtgctgaggggacacaggaatgtgtctgtctgagcaccatgtctat 508
Qy 362 gattggtttgtagctattgtcgctctccatctactacgtatcatcagacagagatc 421
Db 509 gaccgtacagctgcacatcgtcagccatccattacatctgtcattatgacacagctt 568
Qy 422 tgctctcagttgacgactgatactcctggtgttaacgtgtttagtaactcagtggtgtct 481

Db 569 tgcttggtagctatttggtcctctcctgctggygggtctgacaccacagatgtggtctcc 628
Qy 482 accctgactctccagctgcgtcactctgtgaccccatgtgataagatcttctctgtgaa 541
Db 629 acgttccactgtctccacacgctgtgtggaacaatgatgatgacacctctcttctgtgag 688
Qy 542 gtccctgactgtcgaagtattctgtgtgagaacaacagaaatgtagtgaactatc 601
Db 689 atgccccttattgcaactcgtgtgtgtgataccagctcaatgatgatgtgagatgtac 748
Qy 602 ctgttcagttagctcttccatataatcccttgcacatccatccatctatataatgtctt 661
Db 749 ctggcagcttctgtctgtgtgtctcctccttggtggtcctcctctccagccac 808
Qy 662 attgcccagcagttatgagatacaatctcgtcgtgaagatgcagacaaaagatgtggaca 721
Db 809 attgcccggcgtgtgtgaaatcatcagtcacgaagagggaggaagagatcacaacac 868
Qy 722 tgtgttcccatctaaattgtgtgtctcttcttattatagtaagcgtctctgtgactctg 781
Db 869 tgtcttccacagtggtgtgtgtgtctctgtttaaagggagacatcatctcagatctc 928
Qy 782 caaccaccttgcgccagctcccaaggaaggaagatggttctctctctatgaaatc 841
Db 929 cagccagcaagagcaccctccacagcaggaagatgtaagctcgtctctcacaacgta 988
Qy 842 atgacacacatgctgaatcccttataatacattaggaacaagaggtgaag 895
Db 989 gtaactctcgttgaacccattattacaccttgaggagaacccggagtggaag 1042

RESULT 15

US-09-981-566A-27
; Sequence 27, Application US/09981566A
; GENERAL INFORMATION:
; APPLICANT: Kekuda et al.
; TITLE OF INVENTION: Novel GPCR-like Proteins and Nucleic Acids Encoding
; FILE REFERENCE: 21402-163
; CURRENT APPLICATION NUMBER: US/09/981,566A
; CURRENT FILING DATE: 2001-10-16
; PRIOR APPLICATION NUMBER: 60/240,704
; PRIOR FILING DATE: 2000-10-16
; PRIOR APPLICATION NUMBER: 60/262,159
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 60/263,340
; PRIOR FILING DATE: 2001-01-22
; PRIOR APPLICATION NUMBER: 60/264,118
; PRIOR FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: 60/308,203
; PRIOR FILING DATE: 2001-07-27
; PRIOR APPLICATION NUMBER: 60/243,497
; PRIOR FILING DATE: 2000-10-26
; PRIOR APPLICATION NUMBER: 60/244,542
; PRIOR FILING DATE: 2000-10-31
; PRIOR APPLICATION NUMBER: 60/269,031
; PRIOR FILING DATE: 2001-02-15
; PRIOR APPLICATION NUMBER: 60/245,484
; PRIOR FILING DATE: 2000-11-03
; PRIOR APPLICATION NUMBER: 60/255,017
; PRIOR FILING DATE: 2000-12-12
; PRIOR APPLICATION NUMBER: 60/263,216
; PRIOR FILING DATE: 2001-01-22
; PRIOR APPLICATION NUMBER: 60/268,225
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 209
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 27
; LENGTH: 971
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-981-566A-27

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 25, 2002, 19:09:47 ; Search time 217.52 Seconds

(without alignments)
506,482 Million cell updates/sec

Title: US-09-800-321A-4

Perfect score: 1607

Sequence: 1 MMWVDSIIQEFILGFSDR.....NKEVKGFKRLVAVPLIK 313

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 3502263 seqs, 351980561 residues

Total number of hits satisfying chosen parameters: 3502263

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Pending Patents_AA_Main:*
1: /cgn2_6/ptodata/2/paa/PCTUS.COMB.pep.*
2: /cgn2_6/ptodata/2/paa/US06.COMB.pep.*
3: /cgn2_6/ptodata/2/paa/US07.COMB.pep.*
4: /cgn2_6/ptodata/2/paa/US08.COMB.pep.*
5: /cgn2_6/ptodata/2/paa/US081.COMB.pep.*
6: /cgn2_6/ptodata/2/paa/US082.COMB.pep.*
7: /cgn2_6/ptodata/2/paa/US083.COMB.pep.*
8: /cgn2_6/ptodata/2/paa/US084.COMB.pep.*
9: /cgn2_6/ptodata/2/paa/US085.COMB.pep.*
10: /cgn2_6/ptodata/2/paa/US086.COMB.pep.*
11: /cgn2_6/ptodata/2/paa/US087.COMB.pep.*
12: /cgn2_6/ptodata/2/paa/US088.COMB.pep.*
13: /cgn2_6/ptodata/2/paa/US089.COMB.pep.*
14: /cgn2_6/ptodata/2/paa/US090.COMB.pep.*
15: /cgn2_6/ptodata/2/paa/US091.COMB.pep.*
16: /cgn2_6/ptodata/2/paa/US092.COMB.pep.*
17: /cgn2_6/ptodata/2/paa/US093.COMB.pep.*
18: /cgn2_6/ptodata/2/paa/US094.COMB.pep.*
19: /cgn2_6/ptodata/2/paa/US095.COMB.pep.*
20: /cgn2_6/ptodata/2/paa/US096.COMB.pep.*
21: /cgn2_6/ptodata/2/paa/US097.COMB.pep.*
22: /cgn2_6/ptodata/2/paa/US098.COMB.pep.*
23: /cgn2_6/ptodata/2/paa/US099.COMB.pep.*
24: /cgn2_6/ptodata/2/paa/US100.COMB.pep.*
25: /cgn2_6/ptodata/2/paa/US101.COMB.pep.*
26: /cgn2_6/ptodata/2/paa/US60.COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1607	100.0	313	22	US-09-800-321A-4
2	1593	99.1	313	21	US-09-755-017-2
3	1593	99.1	313	21	US-09-795-271-74
4	1593	99.1	313	22	US-09-800-321A-74
5	1593	99.1	313	22	US-09-800-321A-34
6	1593	99.1	313	22	US-09-809-476-74
7	1593	99.1	313	23	US-09-924-359-232

8	1586	98.7	313	22	US-09-800-321A-6	Sequence 6, Appl1
9	1299	80.8	357	21	US-09-795-271-75	Sequence 75, Appl1
10	1299	80.8	357	22	US-09-800-321A-35	Sequence 35, Appl1
11	1299	80.8	357	22	US-09-809-476-75	Sequence 75, Appl1
12	1299	80.8	357	23	US-09-912-976-57	Sequence 57, Appl1
13	1299	80.8	357	23	US-09-912-976-63	Sequence 63, Appl1
14	1295	80.6	310	21	US-09-795-271-77	Sequence 77, Appl1
15	1295	80.6	310	22	US-09-800-321A-36	Sequence 36, Appl1
16	1295	80.6	310	22	US-09-809-476-77	Sequence 77, Appl1
17	1295	80.6	357	21	US-09-795-271-76	Sequence 76, Appl1
18	1295	80.6	357	22	US-09-800-321A-37	Sequence 37, Appl1
19	1295	80.6	357	22	US-09-809-476-76	Sequence 76, Appl1
20	1295	80.6	357	23	US-09-912-976-58	Sequence 58, Appl1
21	1295	80.6	357	23	US-09-912-976-65	Sequence 65, Appl1
22	1295	80.6	357	23	US-09-924-359-142	Sequence 142, Appl1
23	1259	78.3	313	21	US-09-779-679-62	Sequence 62, Appl1
24	1259	78.3	313	21	US-09-795-271-73	Sequence 73, Appl1
25	1259	78.3	313	22	US-09-809-476-73	Sequence 73, Appl1
26	1259	78.3	313	23	US-09-912-976-62	Sequence 62, Appl1
27	1259	78.3	313	23	US-09-924-359-143	Sequence 143, Appl1
28	1247	77.6	313	21	US-09-779-679-55	Sequence 55, Appl1
29	1139	70.9	280	22	US-09-800-321A-38	Sequence 38, Appl1
30	1123	69.9	504	26	US-60-248-505-940	Sequence 940, Appl
31	1034	64.3	203	22	US-09-800-321A-41	Sequence 41, Appl
32	1003	62.4	199	26	US-60-178-308-2276	Sequence 2276, Ap
33	953	59.3	692	26	US-60-258-250-110	Sequence 110, App
34	948	59.0	440	26	US-60-245-225-321	Sequence 321, App
35	948	59.0	440	26	US-60-258-250-97	Sequence 97, Appl
36	944.5	58.8	327	21	US-09-795-271-29	Sequence 29, Appl
37	944.5	58.8	327	22	US-09-809-476-29	Sequence 29, Appl
38	939	58.4	310	19	US-09-524-730-2	Sequence 2, Appl1
39	939	58.4	310	24	US-10-028-521-2	Sequence 2, Appl1
40	937	58.3	312	23	US-09-924-359-2	Sequence 2, Appl1
41	937	58.3	312	23	US-09-924-359-4	Sequence 4, Appl1
42	936	58.2	309	23	US-09-912-976-16	Sequence 16, Appl
43	936	58.2	309	24	US-10-034-842-7	Sequence 7, Appl
44	936	58.2	309	26	US-60-207-360-335	Sequence 335, App
45	936	58.2	309	26	US-60-301-751-12	Sequence 12, Appl

ALIGNMENTS

RESULT 1
US-09-800-321A-4
: Sequence 4, Application US/09800321A
: GENERAL INFORMATION:
: APPLICANT: Padigar, Muralihara
: APPLICANT: Burgess, Catherine E
: APPLICANT: Mishra, Vishnu
: APPLICANT: Li, Li
: APPLICANT: Baumgaertner, Jason C
: APPLICANT: Majumder, Kundu
: APPLICANT: Spytek, Kimberly A
: APPLICANT: Tchiernev, Velizar T
: TITLE OF INVENTION: Novel Proteins and Nucleic Acids Encoding Same
: FILE REFERENCE: 15966-703 US
: CURRENT APPLICATION NUMBER: US/09/800,321A
: PRIOR FILING DATE: 2001-03-05
: PRIOR APPLICATION NUMBER: 60/186,606
: PRIOR FILING DATE: 2000-03-03
: PRIOR APPLICATION NUMBER: 60/221,942
: PRIOR FILING DATE: 2000-07-31
: PRIOR APPLICATION NUMBER: 60/260,285
: PRIOR FILING DATE: 2001-01-08
: PRIOR APPLICATION NUMBER: 60/220,263
: PRIOR FILING DATE: 2000-07-24
: PRIOR APPLICATION NUMBER: 60/257,600
: PRIOR FILING DATE: 2000-12-21
: PRIOR APPLICATION NUMBER: 60/187,295
: PRIOR FILING DATE: 2000-03-06
: PRIOR APPLICATION NUMBER: 60/219,854
: PRIOR FILING DATE: 2000-07-21

```
; PRIOR APPLICATION NUMBER: 60/187,249
; PRIOR FILING DATE: 2000-03-06
; PRIOR APPLICATION NUMBER: 60/187,247
; PRIOR FILING DATE: 2000-03-06
; PRIOR APPLICATION NUMBER: 60/187,250
; PRIOR FILING DATE: 2000-03-06
; PRIOR APPLICATION NUMBER: 60/187,253
; PRIOR FILING DATE: 2000-03-06
; PRIOR APPLICATION NUMBER: 60/187,248
; PRIOR FILING DATE: 2000-03-06
; PRIOR APPLICATION NUMBER: 60/187,296
; PRIOR FILING DATE: 2000-03-06
; PRIOR APPLICATION NUMBER: 60/187,563
; PRIOR FILING DATE: 2000-03-07
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 313
; TYPE: PR
; ORGANISM: Homo sapiens
US-09-800-321a-4
```

```
Query Match          100.0%; Score 1607; DB 22; Length 313;
Best Local Similarity 100.0%; Pred. No. 5.1e-145;
Matches 313; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNWVNDSTIOEFIGLLGFSRPMLEPPLVPLISYTVIFGNLTIIIVSRDLTKLHPMY 60
Db 1 MNWVNDSTIOEFIGLLGFSRPMLEPPLVPLISYTVIFGNLTIIIVSRDLTKLHPMY 60

QY 61 FFLNLSLDDCYTTCYCPQMLVNLCSIRKVIISYRGCAOQFIFALGATEYLLAVNSF 120
Db 61 FFLNLSLDDCYTTCYCPQMLVNLCSIRKVIISYRGCAOQFIFALGATEYLLAVNSF 120

QY 121 DMFAICRPLHYSYMORLCLOLAASWYGFNSVWLSTLTQLPLCDPYVIDHFLCE 180
Db 121 DMFAICRPLHYSYMORLCLOLAASWYGFNSVWLSTLTQLPLCDPYVIDHFLCE 180

QY 181 VPALKISCVEETTANAEFLVSELFHLIPLTLIIISYAFIVRAVLRIOSEGRQKAFGT 240
Db 181 VPALKISCVEETTANAEFLVSELFHLIPLTLIIISYAFIVRAVLRIOSEGRQKAFGT 240

QY 241 CGSHLIIVSLEYSTAVSVYIQPPSPSSKDGKMYSLFYGIAPMLNPLIYTLRNKEVEG 300
Db 241 CGSHLIIVSLEYSTAVSVYIQPPSPSSKDGKMYSLFYGIAPMLNPLIYTLRNKEVEG 300

QY 301 FKRLVAVFLIKK 313
Db 301 FKRLVAVFLIKK 313
```

```
RESULT 2
US-09-755-017-2
; Sequence 2, Application US/09755017
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Wilganowski, Nathaniel
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Friedrich, Glenn
; APPLICANT: Abula, Alejandro
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: Novel Human Membrane Proteins and
; FILE REFERENCE: LEX-0115-USA
; CURRENT APPLICATION NUMBER: US/09/755,017
; PRIOR FILING DATE: 2001-01-05
; PRIOR APPLICATION NUMBER: US 60/175,764
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
```

```
; LENGTH: 313
; TYPE: PR
; ORGANISM: Homo Sapien
US-09-755-017-2
```

```
Query Match          99.1%; Score 1593; DB 21; Length 313;
Best Local Similarity 99.7%; Pred. No. 1.1e-143;
Matches 312; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MNWVNDSTIOEFIGLLGFSRPMLEPPLVPLISYTVIFGNLTIIIVSRDLTKLHPMY 60
Db 1 MNWVNDSTIOEFIGLLGFSRPMLEPPLVPLISYTVIFGNLTIIIVSRDLTKLHPMY 60

QY 61 FFLNLSLDDCYTTCYCPQMLVNLCSIRKVIISYRGCAOQFIFALGATEYLLAVNSF 120
Db 61 FFLNLSLDDCYTTCYCPQMLVNLCSIRKVIISYRGCAOQFIFALGATEYLLAVNSF 120

QY 121 DMFAICRPLHYSYMORLCLOLAASWYGFNSVWLSTLTQLPLCDPYVIDHFLCE 180
Db 121 DMFAICRPLHYSYMORLCLOLAASWYGFNSVWLSTLTQLPLCDPYVIDHFLCE 180

QY 181 VPALKISCVEETTANAEFLVSELFHLIPLTLIIISYAFIVRAVLRIOSEGRQKAFGT 240
Db 181 VPALKISCVEETTANAEFLVSELFHLIPLTLIIISYAFIVRAVLRIOSEGRQKAFGT 240

QY 241 CGSHLIIVSLEYSTAVSVYIQPPSPSSKDGKMYSLFYGIAPMLNPLIYTLRNKEVEG 300
Db 241 CGSHLIIVSLEYSTAVSVYIQPPSPSSKDGKMYSLFYGIAPMLNPLIYTLRNKEVEG 300

QY 301 FKRLVAVFLIKK 313
Db 301 FKRLVAVFLIKK 313
```

```
RESULT 3
US-09-795-271-74
; Sequence 74, Application US/09795271
; GENERAL INFORMATION:
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Majumder, Kumud
; APPLICANT: Burgess, Catherine E
; APPLICANT: Vernet, Corine A.M
; APPLICANT: Fernandes, Elma
; APPLICANT: Shmukets, Richard A
; APPLICANT: Tcherenev, Velizar T
; APPLICANT: Mishra, Vishnu
; APPLICANT: Casman, Stacie
; APPLICANT: Splet, Kimberly A
; APPLICANT: Zehusen, Bryan
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 15966-694
; CURRENT APPLICATION NUMBER: US/09/795,271
; PRIOR FILING DATE: 2001-02-27
; PRIOR APPLICATION NUMBER: 60/185,674
; PRIOR FILING DATE: 2000-02-28
; PRIOR APPLICATION NUMBER: 60/185,535
; PRIOR FILING DATE: 2000-02-28
; PRIOR APPLICATION NUMBER: 60/186,585
; PRIOR FILING DATE: 2000-03-03
; PRIOR APPLICATION NUMBER: 60/186,604
; PRIOR FILING DATE: 2000-03-03
; PRIOR APPLICATION NUMBER: 60/186,584
; PRIOR FILING DATE: 2000-03-03
; PRIOR APPLICATION NUMBER: 60/186,717
; PRIOR FILING DATE: 2000-03-03
; PRIOR APPLICATION NUMBER: 60/186,716
; PRIOR FILING DATE: 2000-03-03
; PRIOR APPLICATION NUMBER: 60/186,719
; PRIOR FILING DATE: 2000-03-03
; PRIOR APPLICATION NUMBER: 60/186,827
; PRIOR FILING DATE: 2000-03-03
; PRIOR APPLICATION NUMBER: 60/218,323
```

```
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/218,435
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/220,517
; PRIOR FILING DATE: 2000-07-24
; PRIOR APPLICATION NUMBER: 60/223,897
; PRIOR FILING DATE: 2000-08-09
; PRIOR APPLICATION NUMBER: 60/260,020
; PRIOR FILING DATE: 2001-01-05
; PRIOR APPLICATION NUMBER: 60/264,849
; PRIOR FILING DATE: 2001-01-29
; PRIOR APPLICATION NUMBER: 60/186,715
; PRIOR FILING DATE: 2000-03-03
; NUMBER OF SEQ ID NOS: 83
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 74
; LENGTH: 313
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-795-271-74
```

Query Match 99.1%; Score 1593; DB 21; Length 313;

Best Local Similarity 99.7%; Pred. No. 1,1e-143; Indels 0; Gaps 0;

Matches 312; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
QY 1 MNWVNDSTIOEFILLGSDRPMLEFPLLVFLISYTVTFGNTLILVSRDTRKLTHTPMY 60
DB 1 MNWVNDSTIOEFILLGSDRPMLEFPLLVFLISYTVTFGNTLILVSRDTRKLTHTPMY 60
QY 61 FEFLNLSLDDCYTTCVPMQVNLCSIRKVISYRGCAQLFIFLAGATEYLLAVMSF 120
DB 61 FEFLNLSLDDCYTTCVPMQVNLCSIRKVISYRGCAQLFIFLAGATEYLLAVMSF 120
QY 121 DWEFAICRPLHYSYMQRCLQLAASWYTGFSNSWMLSTLTQLDPLCDPYVIDHFLCE 180
DB 121 DWEFAICRPLHYSYMQRCLQLAASWYTGFSNSWMLSTLTQLDPLCDPYVIDHFLCE 180
QY 181 VPALKISCVEETANEAEFLVSEFLPLPLTLISYAFIVRAVLRIQSAEGROKAFGT 240
DB 181 VPALKISCVEETANEAEFLVSEFLPLPLTLISYAFIVRAVLRIQSAEGROKAFGT 240
QY 241 CGSHLIYVSLFYSTAVSVYLOPPSPSSKDGKMSLFYGIAPMLNPLIYTLRNKKEVKEG 300
DB 241 CGSHLIYVSLFYSTAVSVYLOPPSPSSKDGKMSLFYGIAPMLNPLIYTLRNKKEVKEG 300
QY 301 FKRLVARVFLIKK 313
DB 301 FKRLVARVFLIKK 313
```

```
RESULT 4
US-09-800-321A-2
; Sequence 2, Application US/09800321A
; GENERAL INFORMATION:
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Burgess, Catherine E
; APPLICANT: Mishra, Vishnu
; APPLICANT: Li, Li
; APPLICANT: Baumgartner, Jason C
; APPLICANT: Majumder, Kumud
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Tcherenev, Velizar T
; TITLE OF INVENTION: Novel Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 15966-703 US
; CURRENT FILING DATE: 2001-03-05
; PRIOR FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: 60/186,606
; PRIOR FILING DATE: 2000-03-03
; PRIOR APPLICATION NUMBER: 60/221,942
; PRIOR FILING DATE: 2000-07-31
; PRIOR APPLICATION NUMBER: 60/260,285
; PRIOR FILING DATE: 2001-01-08
```

```
; PRIOR APPLICATION NUMBER: 60/220,263
; PRIOR FILING DATE: 2000-07-24
; PRIOR APPLICATION NUMBER: 60/257,600
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/187,295
; PRIOR FILING DATE: 2000-03-06
; PRIOR APPLICATION NUMBER: 60/219,854
; PRIOR FILING DATE: 2000-07-21
; PRIOR APPLICATION NUMBER: 60/187,249
; PRIOR FILING DATE: 2000-03-06
; PRIOR APPLICATION NUMBER: 60/187,247
; PRIOR FILING DATE: 2000-03-06
; PRIOR APPLICATION NUMBER: 60/187,250
; PRIOR FILING DATE: 2000-03-06
; PRIOR APPLICATION NUMBER: 60/187,253
; PRIOR FILING DATE: 2000-03-06
; PRIOR APPLICATION NUMBER: 60/187,248
; PRIOR FILING DATE: 2000-03-06
; PRIOR APPLICATION NUMBER: 60/187,296
; PRIOR FILING DATE: 2000-03-06
; PRIOR APPLICATION NUMBER: 60/187,563
; PRIOR FILING DATE: 2000-03-07
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 2
; LENGTH: 313
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-800-321A-2
```

Query Match 99.1%; Score 1593; DB 22; Length 313;

Best Local Similarity 99.7%; Pred. No. 1,1e-143; Indels 0; Gaps 0;

Matches 312; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
QY 1 MNWVNDSTIOEFILLGSDRPMLEFPLLVFLISYTVTFGNTLILVSRDTRKLTHTPMY 60
DB 1 MNWVNDSTIOEFILLGSDRPMLEFPLLVFLISYTVTFGNTLILVSRDTRKLTHTPMY 60
QY 61 FEFLNLSLDDCYTTCVPMQVNLCSIRKVISYRGCAQLFIFLAGATEYLLAVMSF 120
DB 61 FEFLNLSLDDCYTTCVPMQVNLCSIRKVISYRGCAQLFIFLAGATEYLLAVMSF 120
QY 121 DWEFAICRPLHYSYMQRCLQLAASWYTGFSNSWMLSTLTQLDPLCDPYVIDHFLCE 180
DB 121 DWEFAICRPLHYSYMQRCLQLAASWYTGFSNSWMLSTLTQLDPLCDPYVIDHFLCE 180
QY 181 VPALKISCVEETANEAEFLVSEFLPLPLTLISYAFIVRAVLRIQSAEGROKAFGT 240
DB 181 VPALKISCVEETANEAEFLVSEFLPLPLTLISYAFIVRAVLRIQSAEGROKAFGT 240
QY 241 CGSHLIYVSLFYSTAVSVYLOPPSPSSKDGKMSLFYGIAPMLNPLIYTLRNKKEVKEG 300
DB 241 CGSHLIYVSLFYSTAVSVYLOPPSPSSKDGKMSLFYGIAPMLNPLIYTLRNKKEVKEG 300
QY 301 FKRLVARVFLIKK 313
DB 301 FKRLVARVFLIKK 313
```

```
RESULT 5
US-09-800-321A-34
; Sequence 34, Application US/09800321A
; GENERAL INFORMATION:
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Burgess, Catherine E
; APPLICANT: Mishra, Vishnu
; APPLICANT: Li, Li
; APPLICANT: Baumgartner, Jason C
; APPLICANT: Majumder, Kumud
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Tcherenev, Velizar T
; TITLE OF INVENTION: Novel Proteins and Nucleic Acids Encoding Same
```

```
FILE REFERENCE: 15966-703 US
CURRENT APPLICATION NUMBER: US/09/800,321A
CURRENT FILING DATE: 2001-03-05
PRIOR APPLICATION NUMBER: 60/186,606
PRIOR FILING DATE: 2000-03-03
PRIOR APPLICATION NUMBER: 60/221,942
PRIOR FILING DATE: 2000-07-31
PRIOR APPLICATION NUMBER: 60/260,285
PRIOR FILING DATE: 2001-01-08
PRIOR APPLICATION NUMBER: 60/220,263
PRIOR FILING DATE: 2000-07-24
PRIOR APPLICATION NUMBER: 60/257,600
PRIOR FILING DATE: 2000-12-21
PRIOR APPLICATION NUMBER: 60/187,295
PRIOR FILING DATE: 2000-03-06
PRIOR APPLICATION NUMBER: 60/219,854
PRIOR FILING DATE: 2000-07-21
PRIOR APPLICATION NUMBER: 60/187,249
PRIOR FILING DATE: 2000-03-06
PRIOR APPLICATION NUMBER: 60/187,247
PRIOR FILING DATE: 2000-03-06
PRIOR APPLICATION NUMBER: 60/187,250
PRIOR FILING DATE: 2000-03-06
PRIOR APPLICATION NUMBER: 60/187,253
PRIOR FILING DATE: 2000-03-06
PRIOR APPLICATION NUMBER: 60/187,248
PRIOR FILING DATE: 2000-03-06
PRIOR APPLICATION NUMBER: 60/187,296
PRIOR FILING DATE: 2000-03-06
PRIOR APPLICATION NUMBER: 60/187,563
PRIOR FILING DATE: 2000-03-07
NUMBER OF SEQ ID NOS: 78
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 34
LENGTH: 313
TYPE: PRT
ORGANISM: Homo sapiens
US-09-800-321a-34

Query Match          99.1%; Score 1593; DB 22; Length 313;
Best Local Similarity 99.7%; Pred. No. 1,1e-143;
Matches 312; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MNWVDSIIIOEFILGFSDRPMLEFPLLVFLISYTVTIFGNLTIIIVSRDLTKLHTPMY 60
DB 1 MNWVDSIIIOEFILGFSDRPMLEFPLLVFLISYTVTIFGNLTIIIVSRDLTKLHTPMY 60
QY 61 FELTNLSLDDCYTTCVPMQVLNLCIRKVISYRGCAQLFIFLAGATEYLLAVMSF 120
DB 61 FELTNLSLDDCYTTCVPMQVLNLCIRKVISYRGCAQLFIFLAGATEYLLAVMSF 120
QY 121 DMFVAICRPLHYIYIMQRLCLOLAASWYTGFSNSYWLSTLTQLPDCPYVIDHFLCE 180
DB 121 DMFVAICRPLHYIYIMQRLCLOLAASWYTGFSNSYWLSTLTQLPDCPYVIDHFLCE 180
QY 121 DMFVAICRPLHYIYIMQRLCLOLAASWYTGFSNSYWLSTLTQLPDCPYVIDHFLCE 180
DB 121 DMFVAICRPLHYIYIMQRLCLOLAASWYTGFSNSYWLSTLTQLPDCPYVIDHFLCE 180
QY 181 VPALKLSCEVTANAELELVSELPHLIPLTILISYAFIVRAVLRIOSEGRQKAFGT 240
DB 181 VPALKLSCEVTANAELELVSELPHLIPLTILISYAFIVRAVLRIOSEGRQKAFGT 240
QY 241 CGSHLIIVYSFYSTAVSVYIQPPSPSSKDGKMYSLFYGIAPMLNPLIYTLRKKEVKEG 300
DB 241 CGSHLIIVYSFYSTAVSVYIQPPSPSSKDGKMYSLFYGIAPMLNPLIYTLRKKEVKEG 300
QY 301 FKRLVAVRFLIKK 313
DB 301 FKRLVAVRFLIKK 313

RESULT 6
US-09-800-476-74
; Sequence 74, Application US/09809476
; GENERAL INFORMATION:
```

```
APPLICANT: Padigaru, Muralidhara
APPLICANT: Majumder, Kumud
APPLICANT: Burgess, Catherine E
APPLICANT: Vernet, Corine A.M
APPLICANT: Fernandes, Elma
APPLICANT: Shinkets, Richard A
APPLICANT: Tchernev, Velizar T
APPLICANT: Mishra, Vishnu
APPLICANT: Casman, Stacie
APPLICANT: Spytek, Kimberly A
APPLICANT: Zerhusen, Bryan
TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
FILE REFERENCE: 15966-694CIP
CURRENT APPLICATION NUMBER: US/09/809,476
CURRENT FILING DATE: 2002-03-15
PRIOR APPLICATION NUMBER: 60/185,674
PRIOR FILING DATE: 2000-02-28
PRIOR APPLICATION NUMBER: 60/185,535
PRIOR FILING DATE: 2000-02-28
PRIOR APPLICATION NUMBER: 60/186,585
PRIOR FILING DATE: 2000-03-03
PRIOR APPLICATION NUMBER: 60/186,604
PRIOR FILING DATE: 2000-03-03
PRIOR APPLICATION NUMBER: 60/186,584
PRIOR FILING DATE: 2000-03-03
PRIOR APPLICATION NUMBER: 60/186,717
PRIOR FILING DATE: 2000-03-03
PRIOR APPLICATION NUMBER: 60/186,716
PRIOR FILING DATE: 2000-03-03
PRIOR APPLICATION NUMBER: 60/186,719
PRIOR FILING DATE: 2000-03-03
PRIOR APPLICATION NUMBER: 60/186,827
PRIOR FILING DATE: 2000-03-03
PRIOR APPLICATION NUMBER: 60/218,323
PRIOR FILING DATE: 2000-07-14
PRIOR APPLICATION NUMBER: 60/218,435
PRIOR FILING DATE: 2000-07-14
PRIOR APPLICATION NUMBER: 60/220,517
PRIOR FILING DATE: 2000-07-24
PRIOR APPLICATION NUMBER: 60/223,897
PRIOR FILING DATE: 2000-08-09
PRIOR APPLICATION NUMBER: 60/260,020
PRIOR FILING DATE: 2001-01-05
PRIOR APPLICATION NUMBER: 60/264,849
PRIOR FILING DATE: 2001-01-29
PRIOR APPLICATION NUMBER: 60/186,715
PRIOR FILING DATE: 2000-03-03
PRIOR APPLICATION NUMBER: 09/795,271
PRIOR FILING DATE: 2001-02-27
PRIOR APPLICATION NUMBER: 60/259,031
PRIOR FILING DATE: 2000-12-28
NUMBER OF SEQ ID NOS: 92
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 74
LENGTH: 313
TYPE: PRT
ORGANISM: Homo sapiens
US-09-800-476-74

Query Match          99.1%; Score 1593; DB 22; Length 313;
Best Local Similarity 99.7%; Pred. No. 1,1e-143;
Matches 312; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MNWVDSIIIOEFILGFSDRPMLEFPLLVFLISYTVTIFGNLTIIIVSRDLTKLHTPMY 60
DB 1 MNWVDSIIIOEFILGFSDRPMLEFPLLVFLISYTVTIFGNLTIIIVSRDLTKLHTPMY 60
QY 61 FELTNLSLDDCYTTCVPMQVLNLCIRKVISYRGCAQLFIFLAGATEYLLAVMSF 120
DB 61 FELTNLSLDDCYTTCVPMQVLNLCIRKVISYRGCAQLFIFLAGATEYLLAVMSF 120
QY 121 DMFVAICRPLHYIYIMQRLCLOLAASWYTGFSNSYWLSTLTQLPDCPYVIDHFLCE 180
DB 121 DMFVAICRPLHYIYIMQRLCLOLAASWYTGFSNSYWLSTLTQLPDCPYVIDHFLCE 180
```



```

Db 121 DRVAICRPLHYSYIMHQRCLQLAASWYTGFSNSWMLSTLTQLPLCDPYVIDHFLCE 180
QY 181 VPALLKSCVETTANAELELVSELFHLIPLTLLISYAFIVRAVRIQSAEGROKAFGT 240
Db 181 VPALLKSCVETTANAELELVSELFHLIPLTLLISYAFIVRAVRIQSAEGROKAFGT 240
QY 241 CGSHLIIVSLFYSTAVSVYLQPPSPSSKDGKMWSLFYGIAPMLNPLIYTLRNKKEVEG 300
Db 241 CGSHLIIVSLFYSTAVSVYLQPPSPSSKDGKMWSLFYGIAPMLNPLIYTLRNKKEVEG 300
QY 301 FKRLVARVFLIKK 313
Db 301 FKRLVARVFLIKK 313

```

```

RESULT 7
US-09-924-359-232
; Sequence 232, Application US/09924359
; GENERAL INFORMATION:
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Padigar, Muralidhara
; APPLICANT: Zehusen, Bryan D
; APPLICANT: Baumgartner, Jason
; APPLICANT: Li, Li
; APPLICANT: Casman, Stacie
; APPLICANT: Vernet, Corine
; APPLICANT: Ballinger, Robert A
; APPLICANT: Shenoy, Suresh G
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Burgess, Catherine E
; APPLICANT: Mezes, Peter S
; APPLICANT: Grosse, William M
; APPLICANT: Alsobrook II, John P
; APPLICANT: Gorman, Linda
; APPLICANT: Laroche, William J
; APPLICANT: Taupier Jr, Raymond J
; APPLICANT: Coleman, Steven D
; APPLICANT: Szekeres Jr, Edward S
; FILE REFERENCE: 21402-078
; TITLE OF INVENTION: Novel Proteins And Nucleic Acids Encoding Same
; CURRENT APPLICATION NUMBER: US/09/924,359
; PRIOR FILING DATE: 2001-08-07
; PRIOR APPLICATION NUMBER: 60/223,472
; PRIOR FILING DATE: 2000-08-07
; PRIOR APPLICATION NUMBER: 60/223,138
; PRIOR FILING DATE: 2000-08-07
; PRIOR APPLICATION NUMBER: 60/224,613
; PRIOR FILING DATE: 2000-08-11
; PRIOR APPLICATION NUMBER: 60/224,815
; PRIOR FILING DATE: 2000-08-11
; PRIOR APPLICATION NUMBER: 60/260,003
; PRIOR FILING DATE: 2001-01-05
; PRIOR APPLICATION NUMBER: 60/260,072
; PRIOR FILING DATE: 2001-01-05
; PRIOR APPLICATION NUMBER: 60/260,283
; PRIOR FILING DATE: 2001-01-08
; PRIOR APPLICATION NUMBER: 60/260,450
; PRIOR FILING DATE: 2001-01-09
; PRIOR APPLICATION NUMBER: 60/261,156
; PRIOR FILING DATE: 2001-01-10
; PRIOR APPLICATION NUMBER: 60/263,338
; PRIOR FILING DATE: 2001-01-22
; PRIOR APPLICATION NUMBER: 60/263,434
; PRIOR FILING DATE: 2001-01-23
; PRIOR APPLICATION NUMBER: 60/265,704
; PRIOR FILING DATE: 2001-02-01
; PRIOR APPLICATION NUMBER: 60/269,964
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/274,873
; PRIOR FILING DATE: 2001-03-09
; PRIOR APPLICATION NUMBER: 60/276,406
; PRIOR FILING DATE: 2001-03-15

```

```

; PRIOR APPLICATION NUMBER: 60/287,916
; PRIOR FILING DATE: 2001-05-01
; NUMBER OF SEQ ID NOS: 427
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 232
; LENGTH: 313
; TYPE: PRT
; ORGANISM: Rattus sp.
US-09-924-359-232

```

```

Query Match          99.1%; Score 1593; DB 23; Length 313;
Best Local Similarity 99.7%; Pred. No. 1,1e-143;
Matches 312; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 1 MNWVNDSTIOEFILGSDRPMLEFPLLYVELISYTYTIFGNLTIIIVSRDPTKHPMY 60
Db 1 MNWVNDSTIOEFILGSDRPMLEFPLLYVELISYTYTIFGNLTIIIVSRDPTKHPMY 60
QY 61 FFLTNLSLDLCYTTCTVPQMLVNLCSIRVVISYRGCAQLFPLALGATEYLLAVMSF 120
Db 61 FFLTNLSLDLCYTTCTVPQMLVNLCSIRVVISYRGCAQLFPLALGATEYLLAVMSF 120
QY 121 DMFAICRPLHYSYIMHQRCLQLAASWYTGFSNSWMLSTLTQLPLCDPYVIDHFLCE 180
Db 121 DMFAICRPLHYSYIMHQRCLQLAASWYTGFSNSWMLSTLTQLPLCDPYVIDHFLCE 180
QY 181 VPALLKSCVETTANAELELVSELFHLIPLTLLISYAFIVRAVRIQSAEGROKAFGT 240
Db 181 VPALLKSCVETTANAELELVSELFHLIPLTLLISYAFIVRAVRIQSAEGROKAFGT 240
QY 241 CGSHLIIVSLFYSTAVSVYLQPPSPSSKDGKMWSLFYGIAPMLNPLIYTLRNKKEVEG 300
Db 241 CGSHLIIVSLFYSTAVSVYLQPPSPSSKDGKMWSLFYGIAPMLNPLIYTLRNKKEVEG 300
QY 301 FKRLVARVFLIKK 313
Db 301 FKRLVARVFLIKK 313

```

```

RESULT 8
US-09-800-321A-6
; Sequence 6, Application US/09800321A
; GENERAL INFORMATION:
; APPLICANT: Padigar, Muralidhara
; APPLICANT: Burgess, Catherine E
; APPLICANT: Mishra, Vishnu
; APPLICANT: Li, Li
; APPLICANT: Baumgartner, Jason C
; APPLICANT: Majumder, Kumud
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Tchernev, Velizar T
; TITLE OF INVENTION: Novel Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 15966-703 US
; CURRENT APPLICATION NUMBER: US/09/800,321A
; PRIOR FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: 60/186,606
; PRIOR FILING DATE: 2000-03-03
; PRIOR APPLICATION NUMBER: 60/221,942
; PRIOR FILING DATE: 2000-07-31
; PRIOR APPLICATION NUMBER: 60/260,285
; PRIOR FILING DATE: 2001-01-08
; PRIOR APPLICATION NUMBER: 60/220,263
; PRIOR FILING DATE: 2000-07-24
; PRIOR APPLICATION NUMBER: 60/257,600
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/187,295
; PRIOR FILING DATE: 2000-03-06
; PRIOR APPLICATION NUMBER: 60/219,854
; PRIOR FILING DATE: 2000-07-21
; PRIOR APPLICATION NUMBER: 60/187,249
; PRIOR FILING DATE: 2000-03-06
; PRIOR APPLICATION NUMBER: 60/187,247

```

```
; PRIOR FILING DATE: 2000-03-06
; PRIOR APPLICATION NUMBER: 60/187,250
; PRIOR FILING DATE: 2000-03-06
; PRIOR APPLICATION NUMBER: 60/187,253
; PRIOR FILING DATE: 2000-03-06
; PRIOR APPLICATION NUMBER: 60/187,248
; PRIOR FILING DATE: 2000-03-06
; PRIOR APPLICATION NUMBER: 60/187,296
; PRIOR FILING DATE: 2000-03-06
; PRIOR APPLICATION NUMBER: 60/187,563
; PRIOR FILING DATE: 2000-03-07
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 313
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-800-321a-6
```

```
Query Match          98.7%; Score 1586; DB 22; Length 313;
Best Local Similarity 99.0%; Pred. No. 5,2e-143;
Matches 310; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 1 MNWVNDIIIOFFILGSDRPMLEFPLLVFLISYTVTIFGNLTIIIVSRDLTKLHPTMY 60
D 1 MNWVNDIIIOFFILGSDRPMLEFPLLVFLISYTVTIFGNLTIIIVSRDLTKLHPTMY 60
QY 61 FFLTNLSLDDCYTTCVPMQVNLVCSIRKVIISYGCVAOLFIFLAGATEYLLAVMSF 120
D 61 FFLTNLSLDDCYTTCVPMQVNLVCSIRKVIISYGCVAOLFIFLAGATEYLLAVMSF 120
QY 121 DMFVAICRPLHYSIHQRLCLQLAASWYTGFSNSYMLSTLTQLPDCPYVIDHFCE 180
D 121 DMFVAICRPLHYSIHQRLCLQLAASWYTGFSNSYMLSTLTQLPDCPYVIDHFCE 180
QY 181 VPALKIKSCVETANEAEFLVSEFLHILPILTLISYAFIVRAVLRIQSAEGRKAFT 240
D 181 VPALKIKSCVETANEAEFLVSEFLHILPILTLISYAFIVRAVLRIQSAEGRKAFT 240
QY 241 CGSHLIVSLFYSTAVSVYIQPPSSSKDQKMWSLFYGIAPMLNPLIYTLRNKEVKEG 300
D 241 CGSHLIVSLFYSTAVSVYIQPPSSSKDQKMWSLFYGIAPMLNPLIYTLRNKEVKEG 300
QY 301 FKRLVAVFLIKK 313
D 301 FKRLVAVFLIKK 313

RESULT 9
US-09-795-271-75
; Sequence 75, Application US/09795271
; GENERAL INFORMATION:
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Padigaru, Kumud
; APPLICANT: Burgess, Catherine E
; APPLICANT: Vernet, Corine A.M
; APPLICANT: Fernandes, Elma
; APPLICANT: Shinkets, Richard A
; APPLICANT: Tchernev, Velizar T
; APPLICANT: Mishra, Vishnu
; APPLICANT: Casman, Stacie
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Zerhusen, Bryan
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 15966-694
; CURRENT APPLICATION NUMBER: US/09/795,271
; PRIOR FILING DATE: 2001-02-27
; PRIOR APPLICATION NUMBER: 60/185,674
; PRIOR FILING DATE: 2000-02-28
; PRIOR APPLICATION NUMBER: 60/185,535
; PRIOR FILING DATE: 2000-02-28
; PRIOR APPLICATION NUMBER: 60/186,585
```

```
; PRIOR FILING DATE: 2000-03-03
; PRIOR APPLICATION NUMBER: 60/186,604
; PRIOR FILING DATE: 2000-03-03
; PRIOR APPLICATION NUMBER: 60/186,584
; PRIOR FILING DATE: 2000-03-03
; PRIOR APPLICATION NUMBER: 60/186,717
; PRIOR FILING DATE: 2000-03-03
; PRIOR APPLICATION NUMBER: 60/186,716
; PRIOR FILING DATE: 2000-03-03
; PRIOR APPLICATION NUMBER: 60/186,719
; PRIOR FILING DATE: 2000-03-03
; PRIOR APPLICATION NUMBER: 60/186,827
; PRIOR FILING DATE: 2000-03-03
; PRIOR APPLICATION NUMBER: 60/218,323
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/218,435
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/220,517
; PRIOR FILING DATE: 2000-07-24
; PRIOR APPLICATION NUMBER: 60/223,897
; PRIOR FILING DATE: 2000-08-09
; PRIOR APPLICATION NUMBER: 60/260,020
; PRIOR FILING DATE: 2001-01-05
; PRIOR APPLICATION NUMBER: 60/264,849
; PRIOR FILING DATE: 2001-01-29
; PRIOR APPLICATION NUMBER: 60/186,715
; PRIOR FILING DATE: 2000-03-03
; NUMBER OF SEQ ID NOS: 83
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 75
; LENGTH: 357
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-795-271-75
```

```
Query Match          80.8%; Score 1299; DB 21; Length 357;
Best Local Similarity 81.6%; Pred. No. 1,8e-115;
Matches 253; Conservative 24; Mismatches 33; Indels 0; Gaps 0;
```

```
QY 1 MNWVNDIIIOFFILGSDRPMLEFPLLVFLISYTVTIFGNLTIIIVSRDLTKLHPTMY 60
D 1 MNWVNDIIIOFFILGSDRPMLEFPLLVFLISYTVTIFGNLTIIIVSRDLTKLHPTMY 60
QY 61 FFLTNLSLDDCYTTCVPMQVNLVCSIRKVIISYGCVAOLFIFLAGATEYLLAVMSF 120
D 61 FFLTNLSLDDCYTTCVPMQVNLVCSIRKVIISYGCVAOLFIFLAGATEYLLAVMSF 120
QY 121 DMFVAICRPLHYSIHQRLCLQLAASWYTGFSNSYMLSTLTQLPDCPYVIDHFCE 180
D 121 DMFVAICRPLHYSIHQRLCLQLAASWYTGFSNSYMLSTLTQLPDCPYVIDHFCE 180
QY 181 VPALKIKSCVETANEAEFLVSEFLHILPILTLISYAFIVRAVLRIQSAEGRKAFT 240
D 181 VPALKIKSCVETANEAEFLVSEFLHILPILTLISYAFIVRAVLRIQSAEGRKAFT 240
QY 241 CGSHLIVSLFYSTAVSVYIQPPSSSKDQKMWSLFYGIAPMLNPLIYTLRNKEVKEG 300
D 241 CGSHLIVSLFYSTAVSVYIQPPSSSKDQKMWSLFYGIAPMLNPLIYTLRNKEVKEG 300
QY 301 FKRLVAVFL 310
D 301 FKRLVAVFL 310
```

```
RESULT 10
US-09-800-321a-35
; Sequence 35, Application US/09800321A
; GENERAL INFORMATION:
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Padigaru, Burgess, Catherine E
; APPLICANT: Mishra, Vishnu
; APPLICANT: Li, Li
```

```

; APPLICANT: Baumgartner, Jason C
; APPLICANT: Majumder, Kumud
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Tchernev, Velizar T
; TITLE OF INVENTION: Novel Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 15966-703 US
; CURRENT APPLICATION NUMBER: US/09/800,321A
; CURRENT FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: 60/186,606
; PRIOR FILING DATE: 2000-03-03
; PRIOR APPLICATION NUMBER: 60/221,942
; PRIOR FILING DATE: 2000-07-31
; PRIOR APPLICATION NUMBER: 60/260,285
; PRIOR FILING DATE: 2001-01-08
; PRIOR APPLICATION NUMBER: 60/220,263
; PRIOR FILING DATE: 2000-07-24
; PRIOR APPLICATION NUMBER: 60/257,600
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/187,295
; PRIOR FILING DATE: 2000-03-06
; PRIOR APPLICATION NUMBER: 60/219,854
; PRIOR FILING DATE: 2000-07-21
; PRIOR APPLICATION NUMBER: 60/187,249
; PRIOR FILING DATE: 2000-03-06
; PRIOR APPLICATION NUMBER: 60/187,247
; PRIOR FILING DATE: 2000-03-06
; PRIOR APPLICATION NUMBER: 60/187,250
; PRIOR FILING DATE: 2000-03-06
; PRIOR APPLICATION NUMBER: 60/187,253
; PRIOR FILING DATE: 2000-03-06
; PRIOR APPLICATION NUMBER: 60/187,248
; PRIOR FILING DATE: 2000-03-06
; PRIOR APPLICATION NUMBER: 60/187,296
; PRIOR FILING DATE: 2000-03-06
; PRIOR APPLICATION NUMBER: 60/187,563
; PRIOR FILING DATE: 2000-03-07
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 35
; LENGTH: 357
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-800-321A-35

Query Match      80.8%; Score 1299; DB 22; Length 357;
Best Local Similarity 81.6%; Pred. No. 1.8e-115;
Matches 253; Conservative 24; Mismatches 33; Indels 0; Gaps 0;

QY 1 MNNVNDIIIOEFLILGSDRPMLEFPLVFLISYTVTIFGNLTIIIVSRDLTKLHTPMY 60
    ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| :
DB 1 MNNVNSVPEFILLVSDPQWLEIPFVWFLEFSYITITIGNLTIIIVSHVDEKRLHTPMY 60
    ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| :
QY 61 FFLTNLSLDLCYTCGVPQMLNLCIRKIVSYRGCAOLFIALGATEYLLAVNSF 120
    ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| :
DB 61 FFLSNLSLDLCYTCGVPQMLNLCIRKIVSYRGCAOLFIALGATEYLLAVNSF 120
    ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| :
QY 121 DMEVAICRPLHYSVMHQRCLQLAAASWYTGFSNSVMTSLTLQPLCDPYVIDHFLCE 180
    ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| :
DB 121 DMEVAICRPLHYSIMHQRCLQLAAASWYTGFSNSVMTSLTLQPLCDPYVIDHFLCE 180
    ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| :
QY 181 VPLLKISCYETIANEAELEFLVSELEHLPLTLISYAFIVRAVLRIGSAEGRQKAFT 240
    ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| :
DB 181 VPLLKISCYETIANEAELEFLVSELEHLPLTLISYAFIVRAVLRIGSAEGRQKAFT 240
    ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| :
QY 241 CGSHLIYVSLEFYTAASVYVLPSPSSKDGKNVSLFYGIAPLNPILYTLRKKEVKEG 300
    ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| :
DB 241 CGSHLIYVSLEFYTAASVYVLPSPSSKDGKNVSLFYGIAPLNPILYTLRKKEVKEA 300
    ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| :
QY 301 FKRLVARVFL 310
    ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| :
DB 301 FKRLVARVFL 310
    ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| :
```

```

RESULT 11
US-09-809-476-75
; Sequence 75, Application US/09809476
; GENERAL INFORMATION:
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Majumder, Kumud
; APPLICANT: Burgess, Catherine E
; APPLICANT: Vernet, Corine A.M
; APPLICANT: Fernandes, Elma
; APPLICANT: Shinkets, Richard A
; APPLICANT: Tchernev, Velizar T
; APPLICANT: Mishra, Vishnu
; APPLICANT: Casman, Stacie
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Zeihusen, Bryan
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 15966-694CIP
; CURRENT APPLICATION NUMBER: US/09/809,476
; CURRENT FILING DATE: 2002-03-15
; PRIOR APPLICATION NUMBER: 60/185,674
; PRIOR FILING DATE: 2000-02-28
; PRIOR APPLICATION NUMBER: 60/185,535
; PRIOR FILING DATE: 2000-02-28
; PRIOR APPLICATION NUMBER: 60/186,585
; PRIOR FILING DATE: 2000-03-03
; PRIOR APPLICATION NUMBER: 60/186,604
; PRIOR FILING DATE: 2000-03-03
; PRIOR APPLICATION NUMBER: 60/186,584
; PRIOR FILING DATE: 2000-03-03
; PRIOR APPLICATION NUMBER: 60/186,717
; PRIOR FILING DATE: 2000-03-03
; PRIOR APPLICATION NUMBER: 60/186,716
; PRIOR FILING DATE: 2000-03-03
; PRIOR APPLICATION NUMBER: 60/186,719
; PRIOR FILING DATE: 2000-03-03
; PRIOR APPLICATION NUMBER: 60/186,827
; PRIOR FILING DATE: 2000-03-03
; PRIOR APPLICATION NUMBER: 60/218,323
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/218,435
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/220,517
; PRIOR FILING DATE: 2000-07-24
; PRIOR APPLICATION NUMBER: 60/223,897
; PRIOR FILING DATE: 2000-08-09
; PRIOR APPLICATION NUMBER: 60/260,020
; PRIOR FILING DATE: 2001-01-05
; PRIOR APPLICATION NUMBER: 60/264,849
; PRIOR FILING DATE: 2001-01-29
; PRIOR APPLICATION NUMBER: 60/186,715
; PRIOR FILING DATE: 2000-03-03
; PRIOR APPLICATION NUMBER: 09/795,271
; PRIOR FILING DATE: 2001-02-27
; PRIOR APPLICATION NUMBER: 60/259,031
; PRIOR FILING DATE: 2000-12-28
; NUMBER OF SEQ ID NOS: 92
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 75
; LENGTH: 357
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-809-476-75

Query Match      80.8%; Score 1299; DB 22; Length 357;
Best Local Similarity 81.6%; Pred. No. 1.8e-115;
Matches 253; Conservative 24; Mismatches 33; Indels 0; Gaps 0;

QY 1 MNNVNDIIIOEFLILGSDRPMLEFPLVFLISYTVTIFGNLTIIIVSRDLTKLHTPMY 60
    ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| :
DB 1 MNNVNSVPEFILLVSDPQWLEIPFVWFLEFSYITITIGNLTIIIVSHVDEKRLHTPMY 60
    ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| :
```

Qy	61	FEENLSTLDDCYTTCYVPMQVLWCSIRKVIYSRGVADLFEFLAGATEVILLVMSF	120
Db	61	FFLSLSTLDDCYTTSYVPMQVLWICNTRKVIISGGVADLFEFLAGSTECILLVMSF	120
Qy	121	DMFVAICRPLHYSVIMHQRLCLQALAAASWYTGFSNSVMSLTLLQLDPCRYVIDHFLCE	180
Db	121	DRFVAICRPLHYSIIMHQRLCFQLAAASWISGFSNSVLSQTMWTLKMLDCHKEVDHFCE	180
Qy	181	VPALIKISCVEETNEAELEFVSLFHLPLTLLISYAFIVRVLVLSQAEGRQRAFGT	240
Db	181	VPALKIKSCVPTNMEAELEFISVLFLLIPVTLTLLISYAFIVQVNLVLSQAEGRRAFGT	240
Qy	241	CGSHLIIVSLFYSYAVSIVLQPPSPSSKDDGKMVSLFYGIAPALNPLIYTLRNKEVKEG	300
Db	241	CGSHLIIVSLFYSGAIVSMYLDQPPSSSKDDGKMVSLFCGIAPALNPLIYTLRNKEVKEA	300
Qy	301	EKKRLVARYFL 310	
Db	301	EKKRLVAKSLL 310	

```

RESULT 12
US-09-912-976-57
; Sequence 57, Application US/09912976
; GENERAL INFORMATION:
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Mezes, Peter
; APPLICANT: Burgess, Catherine
; APPLICANT: Casman, Stacie
; APPLICANT: Groose, William M
; APPLICANT: Alsodrook II, John P
; APPLICANT: Lepley, Denise M
; APPLICANT: Gerlach, Valerie L
; APPLICANT: Macdougall, John R
; APPLICANT: Smithson, Glenda
; APPLICANT: Mishra, Vishnu
; TITLE OR INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-070
; CURRENT APPLICATION NUMBER: US/09/912,976
; CURRENT FILING DATE: 2001-07-05
; PRIOR APPLICATION NUMBER: 60/221,336
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 60/238,333
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: 60/260,675
; PRIOR FILING DATE: 2001-01-10
; PRIOR APPLICATION NUMBER: 60/271,025
; PRIOR FILING DATE: 2001-02-22
; PRIOR APPLICATION NUMBER: 60/278,164
; PRIOR FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: 60/280,876
; PRIOR FILING DATE: 2001-04-02
; NUMBER OF SEQ ID NOS: 99
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 57
; LENGTH: 357
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-912-976-57

```

Query Match	80.88;	Score 1299;	DB 23;	Length 357;
Best Local Similarity	81.68;	Pred. No. 1.8e-115;		
Matches 253; Conservative	24;	Mismatches 33;	Indels 0;	Gaps 0;

Qy	Db	Qy
1	1	61
MMWVNDSTIOEIIIGSDRPMLEPFLVAYITATYTGNTLITIIIVSLDRKLRHPM	MMWVNDSTIOEIIIGSDRPMLEPFLVAYITATYTGNTLITIIIVSLDRKLRHPM	FFFLNLSLLDDCYTTCTVPMQNLVNLGSIKRYISTRGCAOLFPLAIGATEYLLAVMSF
60	60	120
61	61	61
FFFLNLSLLDDCYTTCTVPMQNLVNLGSIKRYISTRGCAOLFPLAIGATEYLLAVMSF	FFFLNLSLLDDCYTTCTVPMQNLVNLGSIKRYISTRGCAOLFPLAIGATEYLLAVMSF	FFFLNLSLLDDCYTTCTVPMQNLVNLGSIKRYISTRGCAOLFPLAIGATEYLLAVMSF
120	120	120

QY	121	DMFVACBPLHXSVMHQRLOLOLAASWNGEJNSWLSITLQLOPLDIPYDHDLCE	160
Db	121	DFAVACBPLHXSIIIMHORLFCOLAASWISGFSNSVLOSTWILKMPCCGKREVDHFE	180
QY	181	VPALKLSCGVETTANAEELFVSELEHLIPTLLLSIAFIVRAVLRIO\$AEGROKAFGT	240
Db	181	VPALKLSCGVDTTANAEELFVSELEHLIPTLLLSIAFIVQAVLVRIO\$AEGROKAFGT	240
QY	241	CGSHLIIVSLEFYSAVSVYLOPSPBSKDOGKMW\$LEFYGIAPMLNPLIYTLRNKEVEG	3000
Db	241	CGSHLIIVSLEFYGAISWYLOPSPBSKDOGKMW\$LCGIAPMLNPLIYTLRNKEVEKA	3000
QY	301	EKRLVAVRFL 310	
Db	301	EKRLVAKSL 310	

```

RESULT 13
US-09-912-976-63
; Sequence 63, Application US/09912976
; GENERAL INFORMATION:
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Mezes, Peter
; APPLICANT: Burgess, Catherine
; APPLICANT: Caseman, Stacie
; APPLICANT: Groose, William M
; APPLICANT: Alsobrook II, John P
; APPLICANT: Lepley, Denise M
; APPLICANT: Gerlach, Valerie L
; APPLICANT: MacDougall, John R
; APPLICANT: Smithson, Glenda
; APPLICANT: Mishra, Vishnu
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-070
; CURRENT APPLICATION NUMBER: US/09/912,976
; CURRENT FILING DATE: 2001-07-05
; PRIOR APPLICATION NUMBER: 60/221,336
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 60/238,333
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: 60/260,675
; PRIOR FILING DATE: 2001-01-10
; PRIOR APPLICATION NUMBER: 60/271,025
; PRIOR FILING DATE: 2001-02-22
; PRIOR APPLICATION NUMBER: 60/278,164
; PRIOR FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: 60/280,876
; PRIOR FILING DATE: 2001-04-02
; NUMBER OF SEQ ID NOS: 99
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 63
; LENGTH: 357
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-912-976-63

```

Query Match	80.8%	Score 1299	DB 23	Length 357
Best Local Similarity	81.6%	Pred. No. 1.8e-115		
Matches 253	Conservative 24	Mismatches 33	Indels 0	Gaps 0

```
QY 1 MNWVNDSTIOGEFILLGSDRPMLEFPLLYVFLISYTYTFEGNLITLYVSRDRLKHTPMY 60
Db 1 MNWVKSVPDEFILLVFESDQPMLEIPPEFVMEFYSYITLFGNLITLYSVHDEKLTHTPMY 60
QY 61 FFLTULSLDDCYTTCYPPQMLVNLCSRKVIYSRGCAQLFFIALGATHEYLINAMSF 120
Db 61 FFLSLMSLLDDCYTTSYPPQMLVNLICATRKVIYSGGCAQLFFIALGSTECILLAMCF 120
QY 121 DMFAVACIPRLHYSTIIMHORICLOLAAASWYGEFSNSWLSLTLLDLPDYPVYIDHFLCE 180
Db 121 DMFAVACIPRLHYSTIIMHORICLOLAAASWISGFENSVLSQSTWTLKMLCSCHKEDVDFCE 180
```

; ORGANISM: Homo sapiens
US-09-795-271-77

3	PRIOR APPLICATION	NUMBER: 60/186,600
2	PRIOR FILING DATE:	2000-03-03
1	PRIOR APPLICATION	NUMBER: 60/221,944
2	PRIOR FILING DATE:	2000-07-31
1	PRIOR APPLICATION	NUMBER: 60/260,289
2	PRIOR FILING DATE:	2001-01-08
1	PRIOR APPLICATION	NUMBER: 60/220,265
2	PRIOR FILING DATE:	2000-07-24
1	PRIOR APPLICATION	NUMBER: 60/257,600
2	PRIOR FILING DATE:	2000-12-21
1	PRIOR APPLICATION	NUMBER: 60/187,299
2	PRIOR FILING DATE:	2000-03-06
1	PRIOR APPLICATION	NUMBER: 60/219,855
2	PRIOR FILING DATE:	2000-07-21
1	PRIOR APPLICATION	NUMBER: 60/187,249
2	PRIOR FILING DATE:	2000-03-06
1	PRIOR APPLICATION	NUMBER: 60/187,24
2	PRIOR FILING DATE:	2000-03-06
1	PRIOR APPLICATION	NUMBER: 60/187,256
2	PRIOR FILING DATE:	2000-03-06
1	PRIOR APPLICATION	NUMBER: 60/187,255
2	PRIOR FILING DATE:	2000-03-06
1	PRIOR APPLICATION	NUMBER: 60/187,248
2	PRIOR FILING DATE:	2000-03-06
1	PRIOR APPLICATION	NUMBER: 60/187,296
2	PRIOR FILING DATE:	2000-03-06
1	PRIOR APPLICATION	NUMBER: 60/187,566
2	PRIOR FILING DATE:	2000-03-07

```

; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 36
; LENGTH: 310
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-800-321A-36

```

```

Query Match      80.6%; Score 1295; DB 22; Length 310;
Best Local Similarity 81.3%; Pred No. 3.8e-115;
Matches 252; Conservative 25; Mismatches 33; Indels 0; Gaps 0;

```

```

QY 1 MMWVNDIIQEFILGFSDRPMLPEPLAVFLISYVTITFGNLTITLVSRLDTKLTMPY 60
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 1 MMWVNSVPOEFILVFSQPMLEIPPFVFMLEFSYLTITFGNLTITLVSRLDTKLTMPY 60
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 61 FFLTNLSLDLCYTCCTVQMLVNLCSIRKVISYRGCAQLFFFLALGATEYLLAVMSF 120
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 61 FFLSNLSLDLCYTCCTVQMLVNLCSIRKVISYRGCAQLFFFLALGATEYLLAVMSF 120
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 121 DMFVAICRPLHYSVIMHQRICQLAASWVTGFSNSVLTSLTLOLPICDPYVIDHFLCE 180
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 121 DMFVAICRPLHYSVIMHQRICQLAASWVTGFSNSVLTSLTLOLPICDPYVIDHFLCE 180
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 181 VPALLKLSCEVTANEAEELFLVSEFLPLPLILISYAFIVRAVLRIQSAEGQRKAFGT 240
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 181 VPALLKLSCEVTANEAEELFLVSEFLPLPLILISYAFIVRAVLRIQSAEGQRKAFGT 240
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 241 CGSHLIVSLFYSTANVSYLQPPSPSSKDGKWSLFGIITAPMLNPLIYTLRNKEVKEG 300
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 241 CGSHLIVSLFYSTANVSYLQPPSPSSKDGKWSLFGIITAPMLNPLIYTLRNKEVKEG 300
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 301 FKRIVARVFL 310
    |||||:|||||:
DB 301 FKRIVAKSL 310
    |||||:|||||:

```

Search completed: August 25, 2002, 20:12:51
 Job time: 3784 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 25, 2002, 20:07:17 ; Search time 133.9 Seconds

(without alignments)
558.742 Million cell updates/sec

Title: US-09-800-321A-4

Perfect score: 1607
Sequence: 1 MMWVDSIIQEFILGFSDR.....NKEVGEGRVAVRVLK 313

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 73335 segs, 239027266 residues

Total number of hits satisfying chosen parameters: 733335

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents_AA_New:*

- 1: /cgn2_6/ptodata/2/paa/PCT_US06_NEW_COMB.pep:*
- 2: /cgn2_6/ptodata/2/paa/US06_NEW_COMB.pep:*
- 3: /cgn2_6/ptodata/2/paa/US07_NEW_COMB.pep:*
- 4: /cgn2_6/ptodata/2/paa/US08_NEW_COMB.pep:*
- 5: /cgn2_6/ptodata/2/paa/US09_NEW_COMB.pep:*
- 6: /cgn2_6/ptodata/2/paa/US10_NEW_COMB.pep:*
- 7: /cgn2_6/ptodata/2/paa/US60_NEW_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Length	DB	ID	Description
1	1299	80.8	357	5	US-09-965-422-50	Sequence 50, Appl
2	1299	80.8	357	5	US-09-907-218-46	Sequence 46, Appl
3	1295	80.6	310	5	US-09-907-218-49	Sequence 49, Appl
4	1295	80.6	357	5	US-09-965-422-52	Sequence 52, Appl
5	1295	80.6	357	5	US-09-907-218-48	Sequence 48, Appl
6	1295	80.6	357	6	US-10-032-189-112	Sequence 112, Appl
7	1259	78.3	313	5	US-09-965-422-44	Sequence 44, Appl
8	1259	78.3	313	5	US-09-907-218-45	Sequence 45, Appl
9	1259	78.3	313	5	US-09-907-218-79	Sequence 79, Appl
10	1130	70.3	313	5	US-09-981-566A-70	Sequence 70, Appl
11	947	58.9	318	1	PCT-US02-03635-48	Sequence 48, Appl
12	939	58.4	310	5	US-09-546-986A-2	Sequence 2, Appl
13	936	58.2	309	1	PCT-US02-03635-32	Sequence 32, Appl
14	936	58.2	309	5	US-09-965-422-12	Sequence 12, Appl
15	935	58.2	309	6	US-10-032-189-111	Sequence 111, Appl
16	935	58.2	312	6	US-10-032-189-34	Sequence 34, Appl
17	933	58.1	311	5	US-09-981-566A-32	Sequence 32, Appl
18	931	57.9	312	6	US-10-032-106-7	Sequence 7, Appl
19	931	57.9	312	6	US-10-032-189-109	Sequence 109, Appl
20	917	57.1	312	5	US-09-981-566A-30	Sequence 30, Appl
21	915	56.9	312	5	US-09-981-566A-31	Sequence 31, Appl
22	908	56.5	310	5	US-09-965-422-47	Sequence 47, Appl
23	902	56.1	312	5	US-09-965-422-49	Sequence 49, Appl
24	902	56.1	312	5	US-09-907-218-47	Sequence 47, Appl
25	902	56.1	312	5	US-09-907-218-78	Sequence 78, Appl
26	902	56.1	312	5	US-09-981-566A-29	Sequence 29, Appl

27	891	55.4	334	5	US-09-965-422-10	Sequence 10, Appl
28	889	55.3	312	5	US-09-965-422-16	Sequence 16, Appl
29	886	55.1	317	1	PCT-US02-03635-11	Sequence 11, Appl
30	885	55.1	316	5	US-09-965-422-45	Sequence 45, Appl
31	884	55.0	334	5	US-09-965-422-4	Sequence 4, Appl
32	883	54.9	312	6	US-10-002-945-98	Sequence 98, Appl
33	883	54.9	334	5	US-09-965-422-6	Sequence 6, Appl
34	883	54.9	334	5	US-09-965-422-8	Sequence 8, Appl
35	883	54.9	334	5	US-09-546-986A-8	Sequence 8, Appl
36	879	54.7	340	5	US-09-546-986A-6	Sequence 6, Appl
37	876	54.5	320	5	US-09-981-566A-28	Sequence 28, Appl
38	875	54.4	319	1	PCT-US02-09923-39	Sequence 39, Appl
39	871	54.2	320	5	US-09-981-566A-33	Sequence 33, Appl
40	871	54.2	320	6	US-10-032-189-113	Sequence 113, Appl
41	867	54.0	317	5	US-09-965-422-51	Sequence 51, Appl
42	865	53.8	312	5	US-09-965-422-48	Sequence 48, Appl
43	855.5	53.2	311	1	PCT-US02-09923-33	Sequence 33, Appl
44	855.5	53.2	311	5	US-09-907-218-10	Sequence 10, Appl
45	825.5	51.4	311	6	US-10-023-601-58	Sequence 58, Appl

ALIGNMENTS

RESULT 1
US-09-965-422-50
Sequence 50, Application US/09965422
GENERAL INFORMATION:
APPLICANT: Spytek, Kimberly A
APPLICANT: Casman, Stacie
APPLICANT: Padigaru, Muralidhara
APPLICANT: Dickson, Kevin
APPLICANT: Vernet, Corline
APPLICANT: Spaderna, Steven K
APPLICANT: Shenoy, Suresh G
APPLICANT: Gerlach, Valerie
APPLICANT: Ellerman, Karen
APPLICANT: Edinger, Shlomit
APPLICANT: MacDougall, John R
APPLICANT: Smithson, Glenda
APPLICANT: Li, Li
APPLICANT: Malyankar, Urtil M
APPLICANT: Taylor, Sarah
APPLICANT: Gunther, Erik
APPLICANT: Tchiernev, Velizar T
TITLE OF INVENTION: Novel Proteins and Nucleic Acids Encoding Same
FILE REFERENCE: 21401-132
CURRENT APPLICATION NUMBER: US/09/965,422
PRIOR FILING DATE: 2001-09-27, 286
PRIOR APPLICATION NUMBER: 60/236, 284
PRIOR FILING DATE: 2000-09-28
PRIOR APPLICATION NUMBER: 60/237, 581
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/238, 735
PRIOR FILING DATE: 2000-10-06
PRIOR APPLICATION NUMBER: 60/240, 736
PRIOR FILING DATE: 2000-10-16
PRIOR APPLICATION NUMBER: 60/260, 019
PRIOR FILING DATE: 2001-01-05
PRIOR APPLICATION NUMBER: 60/260, 338
PRIOR FILING DATE: 2001-01-08
PRIOR APPLICATION NUMBER: 60/262, 156
PRIOR FILING DATE: 2001-01-17
PRIOR APPLICATION NUMBER: 60/262, 498
PRIOR FILING DATE: 2001-01-18
PRIOR APPLICATION NUMBER: 60/263, 133
PRIOR FILING DATE: 2001-01-19
PRIOR APPLICATION NUMBER: 60/263, 691
PRIOR FILING DATE: 2001-01-24
PRIOR APPLICATION NUMBER: 60/266, 109
PRIOR FILING DATE: 2001-02-02

;; PRIOR APPLICATION NUMBER: 60/271,634
;; PRIOR FILING DATE: 2001-02-26
;; NUMBER OF SEQ ID NOS: 127
;; SOFTWARE: PatentIn Ver. 2.1
;; SEQ ID NO 50
;; LENGTH: 357
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-09-965-422-50

Query Match 80.8%; Score 1299; DB 5; Length 357;
Best Local Similarity 81.6%; Pred. No. 6.1e-96;
Matches 253; Conservative 24; Mismatches 33; Indels 0; Gaps 0;

QY 1 MNWVNDSTIOEFGSDRPMLEFPLLVFLISTYVTFGNLTIIIVSRDLTKLHPMY 60
D 1 MNWVNVSVQGEFILLVSDQPMLEIPFVMTLFSYILTFGNLTIIIVSHVDFKLHPMY 60
QY 61 FFLNLSLDDCYTTCYTPQMLVNLCSIRKVIYSGCYAQLFIFLALGATEYLLAVNSF 120
D 61 FFLNLSLDDCYTTCYTPQMLVNLCSIRKVIYSGCYAQLFIFLALGATEYLLAVNSF 120
QY 121 DMFAICRPLHYSTMORLCLOLAASWVGFSNSVWLSLTITLPLCDPYVIDHFCE 180
D 121 DMFAICRPLHYSTMORLCLOLAASWVGFSNSVWLSLTITLPLCDPYVIDHFCE 180
QY 181 VPALLKSCVETTANAELEFLVSELFLHPLTLILISYAFIVRAVLRIQSAEGRKAFGT 240
D 181 VPALLKSCVETTANAELEFLVSELFLHPLTLILISYAFIVRAVLRIQSAEGRKAFGT 240
QY 241 CGSHLIYVSLEYTAVSVYLOPPSPSSKDGKMSLFYGIAPMLNPLITYLRNKEVEG 300
D 241 CGSHLIYVSLEYTAVSVYLOPPSPSSKDGKMSLFYGIAPMLNPLITYLRNKEVEG 300
QY 301 FKRLVAVFL 310
D 301 FKRLVAVFL 310

RESULT 2
; Sequence 46, Application US/09907218
; GENERAL INFORMATION:
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Mishra, Vishnu
; APPLICANT: Paturaju, Meera
; APPLICANT: Tallon, Bruce
; APPLICANT: Casman, Stacie J
; APPLICANT: Wolenc, Adam Ryan
; APPLICANT: Li, Li
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Spytek, Kimberly Ann
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-061
; CURRENT APPLICATION NUMBER: US/09/907, 218
; CURRENT FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: 60/218,746
; PRIOR FILING DATE: 2000-07-17
; PRIOR APPLICATION NUMBER: 60/260,977
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: 60/263,801
; PRIOR FILING DATE: 2001-01-24
; PRIOR APPLICATION NUMBER: 60/268,226
; PRIOR FILING DATE: 2001-02-12
; PRIOR APPLICATION NUMBER: 60/271,622
; PRIOR FILING DATE: 2001-02-26
; NUMBER OF SEQ ID NOS: 89
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 46
; LENGTH: 357
; TYPE: PRT
; ORGANISM: Homo sapiens

US-09-907-218-46

Query Match 80.8%; Score 1299; DB 5; Length 357;
Best Local Similarity 81.6%; Pred. No. 6.1e-96;
Matches 253; Conservative 24; Mismatches 33; Indels 0; Gaps 0;

QY 1 MNWVNDSTIOEFGSDRPMLEFPLLVFLISTYVTFGNLTIIIVSRDLTKLHPMY 60
D 1 MNWVNVSVQGEFILLVSDQPMLEIPFVMTLFSYILTFGNLTIIIVSHVDFKLHPMY 60
QY 61 FFLNLSLDDCYTTCYTPQMLVNLCSIRKVIYSGCYAQLFIFLALGATEYLLAVNSF 120
D 61 FFLNLSLDDCYTTCYTPQMLVNLCSIRKVIYSGCYAQLFIFLALGATEYLLAVNSF 120
QY 121 DMFAICRPLHYSTMORLCLOLAASWVGFSNSVWLSLTITLPLCDPYVIDHFCE 180
D 121 DMFAICRPLHYSTMORLCLOLAASWVGFSNSVWLSLTITLPLCDPYVIDHFCE 180
QY 181 VPALLKSCVETTANAELEFLVSELFLHPLTLILISYAFIVRAVLRIQSAEGRKAFGT 240
D 181 VPALLKSCVETTANAELEFLVSELFLHPLTLILISYAFIVRAVLRIQSAEGRKAFGT 240
QY 241 CGSHLIYVSLEYTAVSVYLOPPSPSSKDGKMSLFYGIAPMLNPLITYLRNKEVEG 300
D 241 CGSHLIYVSLEYTAVSVYLOPPSPSSKDGKMSLFYGIAPMLNPLITYLRNKEVEG 300
QY 301 FKRLVAVFL 310
D 301 FKRLVAVFL 310

RESULT 3
; Sequence 49, Application US/09907218
; GENERAL INFORMATION:
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Mishra, Vishnu
; APPLICANT: Paturaju, Meera
; APPLICANT: Tallon, Bruce
; APPLICANT: Casman, Stacie J
; APPLICANT: Wolenc, Adam Ryan
; APPLICANT: Li, Li
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Spytek, Kimberly Ann
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-061
; CURRENT APPLICATION NUMBER: US/09/907, 218
; CURRENT FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: 60/218,746
; PRIOR FILING DATE: 2000-07-17
; PRIOR APPLICATION NUMBER: 60/260,977
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: 60/263,801
; PRIOR FILING DATE: 2001-01-24
; PRIOR APPLICATION NUMBER: 60/268,226
; PRIOR FILING DATE: 2001-02-12
; PRIOR APPLICATION NUMBER: 60/271,622
; PRIOR FILING DATE: 2001-02-26
; NUMBER OF SEQ ID NOS: 89
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 49
; LENGTH: 310
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-907-218-49

Query Match 80.8%; Score 1295; DB 5; Length 310;
Best Local Similarity 81.3%; Pred. No. 1.1e-95;
Matches 252; Conservative 25; Mismatches 33; Indels 0; Gaps 0;

QY 1 MNWVNDSTIOEFGSDRPMLEFPLLVFLISTYVTFGNLTIIIVSRDLTKLHPMY 60


```

Db      1 MNWVNSVPOEFLILVSDPWLIEPFWFELSYLTITFGNLTILIVSHVDRLHTPMY 60
QY      61 FFLTNLSLDLCYTTCTVPMVLNLCISIRKVISYRCVAOLFPLALGATEYLLAVMSF 120
        |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db      61 FFLSNLSLDLCYTTSTVPQMLVNICNTRKVISYGCVAOLFPLALGSTECILLAVMCF 120
QY      121 DMFVAICRPLHYSIMHQRICFQIAAASWTSVLSYLTITQPLCDPYVIDHFLCE 180
        |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db      121 DMFVAICRPLHYSIMHQRICFQIAAASWTSVLSYLTITQPLCDPYVIDHFLCE 180
QY      181 VPALLTSCVETANEAELEFVSELFLIPLTILISYAFIVRAVRIQSAEGROKAFGT 240
        |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db      181 VPALLTSCVETANEAELEFVSELFLIPLTILISYAFIVRAVRIQSAEGROKAFGT 240
QY      241 CGSHLIYVSLFYGTAVSYLQPPSPSSKDGKMWSLFYGIAPMLNPLIYTLRNKEVKEG 300
        |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db      241 CGSHLIYVSLFYGTAVSYLQPPSPSSKDGKMWSLFYGIAPMLNPLIYTLRNKEVKEA 300
QY      301 FKRLVAVRFL 310
        |||:|||||:
Db      301 FKRLVAKSL 310

```

RESULT 4

```

US-09-965-422-52
; Sequence 52, Application US/09965422
; GENERAL INFORMATION:
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Casman, Stacie
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Dickson, Kevin
; APPLICANT: Vermet, Corlene
; APPLICANT: Spaderina, Steven K
; APPLICANT: Shenoy, Suresh G
; APPLICANT: Gerlach, Valerie
; APPLICANT: Ellerman, Karen
; APPLICANT: Edinger, Shlomit
; APPLICANT: Macdougall, John R
; APPLICANT: Smithson, Glenda
; APPLICANT: Li, Li
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Taylor, Sarah
; APPLICANT: Gunther, Erik
; APPLICANT: Tcheney, Velizar T
; TITLE OF INVENTION: Novel Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21401-132
; CURRENT APPLICATION NUMBER: US/09/965,422
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: 60/236,286
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: 60/236,284
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: 60/237,581
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/238,735
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/240,736
; PRIOR FILING DATE: 2000-10-16
; PRIOR APPLICATION NUMBER: 60/260,019
; PRIOR FILING DATE: 2001-01-05
; PRIOR APPLICATION NUMBER: 60/260,338
; PRIOR FILING DATE: 2001-01-08
; PRIOR APPLICATION NUMBER: 60/262,156
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 60/262,498
; PRIOR FILING DATE: 2001-01-18
; PRIOR APPLICATION NUMBER: 60/263,133
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: 60/263,691
; PRIOR FILING DATE: 2001-01-24
; PRIOR APPLICATION NUMBER: 60/266,109
; PRIOR FILING DATE: 2001-02-02

```

```

; PRIOR APPLICATION NUMBER: 60/271,634
; PRIOR FILING DATE: 2001-02-26
; NUMBER OF SEQ ID NOS: 127
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 52
; LENGTH: 357
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-965-422-52

```

Query Match 80.6%; Score 1295; DB 5; Length 357;
 Best Local Similarity 81.3%; Pred., No. 1.3e-95;
 Matches 252; Conservative 25; Mismatches 33; Indels 0; Gaps 0;

```

QY      1 MNWVNSIIEQFILLGHSRPMLEFPLLVFLISYTYTIFGNLTILIVSRIDTFLHTPMY 60
        |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db      1 MNWVNSVPOEFLILVSDPWLIEPFWFELSYLTITFGNLTILIVSHVDRLHTPMY 60
QY      61 FFLTNLSLDLCYTTCTVPMVLNLCISIRKVISYRCVAOLFPLALGATEYLLAVMSF 120
        |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db      61 FFLSNLSLDLCYTTSTVPQMLVNICNTRKVISYGCVAOLFPLALGSTECILLAVMCF 120
QY      121 DMFVAICRPLHYSIMHQRICFQIAAASWTSVLSYLTITQPLCDPYVIDHFLCE 180
        |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db      121 DMFVAICRPLHYSIMHQRICFQIAAASWTSVLSYLTITQPLCDPYVIDHFLCE 180
QY      181 VPALLTSCVETANEAELEFVSELFLIPLTILISYAFIVRAVRIQSAEGROKAFGT 240
        |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db      181 VPALLTSCVETANEAELEFVSELFLIPLTILISYAFIVRAVRIQSAEGROKAFGT 240
QY      241 CGSHLIYVSLFYGTAVSYLQPPSPSSKDGKMWSLFYGIAPMLNPLIYTLRNKEVKEG 300
        |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db      241 CGSHLIYVSLFYGTAVSYLQPPSPSSKDGKMWSLFYGIAPMLNPLIYTLRNKEVKEA 300
QY      301 FKRLVAVRFL 310
        |||:|||||:
Db      301 FKRLVAKSL 310

```

RESULT 5

```

US-09-907-218-48
; Sequence 48, Application US/09907218
; GENERAL INFORMATION:
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Mishra, Vishnu
; APPLICANT: Patirajan, Meera
; APPLICANT: Tallon, Bruce
; APPLICANT: Casman, Stacie J
; APPLICANT: Wolenc, Adam Ryan
; APPLICANT: Li, Li
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Spytek, Kimberly Ann
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-061
; CURRENT APPLICATION NUMBER: US/09/907,218
; PRIOR FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: 60/218,746
; PRIOR FILING DATE: 2000-07-17
; PRIOR APPLICATION NUMBER: 60/260,977
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: 60/263,801
; PRIOR FILING DATE: 2001-01-24
; PRIOR APPLICATION NUMBER: 60/268,226
; PRIOR FILING DATE: 2001-02-12
; PRIOR APPLICATION NUMBER: 60/271,622
; PRIOR FILING DATE: 2001-02-26
; NUMBER OF SEQ ID NOS: 89
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 48
; LENGTH: 357
; TYPE: PRT
; ORGANISM: Homo sapiens

```

US-09-907-218-48

Query Match 80.6%; Score 1295; DB 5; Length 357;
Best Local Similarity 81.3%; Pred. No. 1.3e-95;
Matches 252; Conservative 25; Mismatches 33; Indels 0; Gaps 0;

QY 1 MNNVNDSTIOEFILGSDRPMLEFLLVFLSYVTYITFGNTIILVSRDLRKLHPTMY 60
DB 1 MNNVNSVPOEFILVSDPMLEIPFVMEFLRSYITLITFGNTIILVSHVDRKLHPTMY 60
QY 61 FFLTNLSLDLCYTTCTVPQMLVNLCSIRKVISYRGCAQLFTFLAGATEYLLAVMSF 120
DB 61 FFLTNLSLDLCYTTCTVPQMLVNLCSIRKVISYRGCAQLFTFLAGATEYLLAVMSF 120
QY 121 DMFVAICRPLHYSTIMHQRCLQLAASWVTGFSNSVWMLSTLTQLPDCPIYIDHFLCE 180
DB 121 DMFVAICRPLHYSTIMHQRCLQLAASWVTGFSNSVWMLSTLTQLPDCPIYIDHFLCE 180
QY 181 VPALLKLSCVETTANAEELFLVSELPHLIPLTILISYAFIVRAVLRIQSAEGROKAFGT 240
DB 181 VPALLKLSCVETTANAEELFLVSELPHLIPLTILISYAFIVRAVLRIQSAEGROKAFGT 240
QY 241 CGSHLIYVSLFYGTAVSVYIQPPSPSSKDGKNVSLFYGIAPMLNPLIYTLRNKEVKEG 300
DB 241 CGSHLIYVSLFYGTAVSVYIQPPSPSSKDGKNVSLFYGIAPMLNPLIYTLRNKEVKEA 300
QY 301 FKRLVARVFL 310
DB 301 FKRLVARVFL 310

RESULT 6
US-10-032-189-112

; Sequence 112, Application US/10032189
; GENERAL INFORMATION:

; APPLICANT: Alsobrook II, John P
; APPLICANT: Tchernev, Velizar T
; APPLICANT: Liu, Xiaohong
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Zerhusen, Bryan D
; APPLICANT: Patturajan, Meera
; APPLICANT: Grose, William M
; APPLICANT: Lepley, Denise M
; APPLICANT: Burgess, Catherine E
; APPLICANT: Shimkets, Richard A
; APPLICANT: Grose, William M
; APPLICANT: Szekeres, Edward S
; APPLICANT: Vernet, Corine A.M.
; APPLICANT: Li, Li
; APPLICANT: Casman, Stacie J
; APPLICANT: Boldog, Ferenc L
; APPLICANT: Gorman, Linda
; APPLICANT: Gangoli, Esha A
; APPLICANT: Rieger, Daniel K
; APPLICANT: Edinger, Shlomit R
; APPLICANT: Gunther, Erik
; APPLICANT: Millet, Isabelle
; APPLICANT: Sciore, Paul
; APPLICANT: Ellerman, Karen
; APPLICANT: MacDougall, John R
; APPLICANT: Smithson, Glenda
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-228
; CURRENT APPLICATION NUMBER: US/10/032,189
; PRIOR FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 60/257,495
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/258,171
; PRIOR FILING DATE: 2000-12-20
; PRIOR APPLICATION NUMBER: 60/269,940
; PRIOR FILING DATE: 2001-02-20

; PRIOR APPLICATION NUMBER: 60/274,192
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: 60/277,826
; PRIOR FILING DATE: 2001-03-22
; PRIOR APPLICATION NUMBER: 60/279,840
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/282,981
; PRIOR FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 60/283,656
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/309,247
; PRIOR FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: 60/311,754
; PRIOR FILING DATE: 2001-08-17
; PRIOR APPLICATION NUMBER: 60/313,331
; PRIOR FILING DATE: 2001-08-17
; NUMBER OF SEQ ID NOS: 260
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 112
; LENGTH: 357
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-032-189-112

Query Match 80.6%; Score 1295; DB 6; Length 357;
Best Local Similarity 81.3%; Pred. No. 1.3e-95;
Matches 252; Conservative 25; Mismatches 33; Indels 0; Gaps 0;

QY 1 MNNVNDSTIOEFILGSDRPMLEFLLVFLSYVTYITFGNTIILVSRDLRKLHPTMY 60
DB 1 MNNVNSVPOEFILVSDPMLEIPFVMEFLRSYITLITFGNTIILVSHVDRKLHPTMY 60
QY 61 FFLTNLSLDLCYTTCTVPQMLVNLCSIRKVISYRGCAQLFTFLAGATEYLLAVMSF 120
DB 61 FFLTNLSLDLCYTTCTVPQMLVNLCSIRKVISYRGCAQLFTFLAGATEYLLAVMSF 120
QY 121 DMFVAICRPLHYSTIMHQRCLQLAASWVTGFSNSVWMLSTLTQLPDCPIYIDHFLCE 180
DB 121 DMFVAICRPLHYSTIMHQRCLQLAASWVTGFSNSVWMLSTLTQLPDCPIYIDHFLCE 180
QY 181 VPALLKLSCVETTANAEELFLVSELPHLIPLTILISYAFIVRAVLRIQSAEGROKAFGT 240
DB 181 VPALLKLSCVETTANAEELFLVSELPHLIPLTILISYAFIVRAVLRIQSAEGROKAFGT 240
QY 241 CGSHLIYVSLFYGTAVSVYIQPPSPSSKDGKNVSLFYGIAPMLNPLIYTLRNKEVKEG 300
DB 241 CGSHLIYVSLFYGTAVSVYIQPPSPSSKDGKNVSLFYGIAPMLNPLIYTLRNKEVKEA 300
QY 301 FKRLVARVFL 310
DB 301 FKRLVARVFL 310

RESULT 7
US-09-965-422-44

; Sequence 44, Application US/09965422
; GENERAL INFORMATION:

; APPLICANT: Spytek, Kimberly A
; APPLICANT: Casman, Stacie
; APPLICANT: Padigar, Muralidhara
; APPLICANT: Dickson, Kevin
; APPLICANT: Vernet, Corine
; APPLICANT: Spaderina, Steven K
; APPLICANT: Shenoy, Suresh G
; APPLICANT: Gerlach, Valerie
; APPLICANT: Ellerman, Karen
; APPLICANT: Edinger, Shlomit
; APPLICANT: MacDougall, John R
; APPLICANT: Smithson, Glenda
; APPLICANT: Li, Li
; APPLICANT: Malvankar, Urial M
; APPLICANT: Taylor, Sarah

```

? APPLICANT: Gunther, Erik
? APPLICANT: Tchernev, Velizar T
? TITLE OF INVENTION: Novel Proteins and Nucleic Acids Encoding Same
? FILE REFERENCE: 21401-132
? CURRENT APPLICATION NUMBER: US/09/965,422
? CURRENT FILING DATE: 2001-09-27
? PRIOR APPLICATION NUMBER: 60/236,286
? PRIOR FILING DATE: 2000-09-28
? PRIOR APPLICATION NUMBER: 60/236,284
? PRIOR FILING DATE: 2000-09-28
? PRIOR APPLICATION NUMBER: 60/237,581
? PRIOR FILING DATE: 2000-10-03
? PRIOR APPLICATION NUMBER: 60/238,755
? PRIOR FILING DATE: 2000-10-06
? PRIOR APPLICATION NUMBER: 60/240,736
? PRIOR FILING DATE: 2000-10-16
? PRIOR APPLICATION NUMBER: 60/260,019
? PRIOR FILING DATE: 2001-01-05
? PRIOR APPLICATION NUMBER: 60/260,338
? PRIOR FILING DATE: 2001-01-08
? PRIOR APPLICATION NUMBER: 60/262,156
? PRIOR FILING DATE: 2001-01-17
? PRIOR APPLICATION NUMBER: 60/262,498
? PRIOR FILING DATE: 2001-01-18
? PRIOR APPLICATION NUMBER: 60/263,133
? PRIOR FILING DATE: 2001-01-19
? PRIOR APPLICATION NUMBER: 60/263,691
? PRIOR FILING DATE: 2001-01-24
? PRIOR APPLICATION NUMBER: 60/266,109
? PRIOR FILING DATE: 2001-02-02
? PRIOR APPLICATION NUMBER: 60/271,634
? PRIOR FILING DATE: 2001-02-26
? NUMBER OF SEQ ID NOS: 127
? SOFTWARE: PatentIn Ver. 2.1
? SEQ ID NO 44
? LENGTH: 313
? TYPE: PRT
? ORGANISM: Rattus norvegicus
? OS-09-965-422-44

```

```

1  APPLICANT: Kekuda, Ramesh
2  APPLICANT: Spytek, Kimberly Ann
3  TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
4  FILE REFERENCE: 21402-061
5  CURRENT APPLICATION NUMBER: US/09/907,218
6  CURRENT FILING DATE: 2002-04-04
7  PRIOR APPLICATION NUMBER: 60/218,746
8  PRIOR FILING DATE: 2000-07-17
9  PRIOR APPLICATION NUMBER: 60/260,977
10 PRIOR FILING DATE: 2001-01-11
11 PRIOR APPLICATION NUMBER: 60/263,801
12 PRIOR FILING DATE: 2001-01-24
13 PRIOR APPLICATION NUMBER: 60/268,226
14 PRIOR FILING DATE: 2001-02-12
15 PRIOR APPLICATION NUMBER: 60/271,622
16 PRIOR FILING DATE: 2001-02-26
17 NUMBER OF SEQ ID NOS: 89
18 SOFTWARE: PatentIn Ver. 2.1
19 SEQ ID NO 79
20 LENGTH: 313
21 TYPE: prt
22 ORGANISM: Rattus norvegicus
23 US-09-907-218-79

```

Query Match	78.3%	Score 1259;	DB 5;	Length 313;
Best Local Similarity	78.0%	Pred. No. 8.3e-93;		
Matches 244;	Conservative 29;	Mismatches 40;	Indels 0;	Gaps 0

[illegible]

```

RESULT 10
US-09-981-566A-70
Sequence 70: Application US/09981566A
GENERAL INFORMATION:
APPLICANT: kekida et al.
TITLE OF INVENTION: Novel GPCR-like Proteins and Nucleic Acids Encoding
FILE REFERENCE: 21402-163
CURRENT APPLICATION NUMBER: US/09/981,566A
CURRENT FILING DATE: 2001-10-16
PRIOR APPLICATION NUMBER: 60/240,704
PRIOR FILING DATE: 2000-10-16
PRIOR APPLICATION NUMBER: 60/262,159
PRIOR FILING DATE: 2001-01-17
PRIOR APPLICATION NUMBER: 60/263,340
PRIOR FILING DATE: 2001-01-22
PRIOR APPLICATION NUMBER: 60/264,118
PRIOR FILING DATE: 2001-01-25
PRIOR APPLICATION NUMBER: 60/308,203

```

```

1 PRIOR FILING DATE: 2001-07-27
2 PRIOR APPLICATION NUMBER: 60/243,497
3 PRIOR FILING DATE: 2000-10-26
4 PRIOR APPLICATION NUMBER: 60/244,562
5 PRIOR FILING DATE: 2000-10-31
6 PRIOR APPLICATION NUMBER: 60/269,031
7 PRIOR FILING DATE: 2001-02-15
8 PRIOR APPLICATION NUMBER: 60/245,484
9 PRIOR FILING DATE: 2000-11-03
10 PRIOR APPLICATION NUMBER: 60/255,017
11 PRIOR FILING DATE: 2000-12-12
12 PRIOR APPLICATION NUMBER: 60/265,216
13 PRIOR FILING DATE: 2001-01-22
14 PRIOR APPLICATION NUMBER: 60/268,225
15 PRIOR FILING DATE: 2001-02-12
16 NUMBER OF SEQ ID NOS: 209
17 SOFTWARE: PatentIn Ver. 2.1
18 SEQ ID NO 70
19 LENGTH: 313
20 TYPE: prt
21 ORGANISM: Homo sapiens
22 US-09-981-566A-70

```

Query Match	70.3%	Score 1130;	DB 5;	Length 313;
Best Local Similarity	70.3%	Pred. No. 1.6e-82;		
Matches 220;	Conservative 34;	Mismatches 59;	Indels 0;	Gaps 0;

Qy	1	MNVWVDS:IIIOEIIILGFSDBRWLEFEPLVFLVLSYTVIIGCNLLIIVSRDRLKTHPMX	60
Qy	1	MNVWVDS:IIIOEIIILGFSDBRWLEFEPLVFLVLSYTVIIGCNLLIIVSRDRLKTHPMX	60
Db	1	MNMENESSPKREFIILIGFSDBRAMLQMPLEFVLLISYITITIFGVNSIMWVCILDPKRLTHPMX	60
Qy	61	FFLTLNLILDLCTYTCYVPOMLVNLMSIRKXISVRCGVAOLFELALGAREYLLILANMSF	120
Db	61	FFLTLNLILDLCTYTTVYPHMLVINGCKTKTISAGCAVAILHIFLALGARECCLILANMSF	120
Qy	121	DMFEAICRPHYSVIMORLOLQALAAASWYVGFSGNSVNLSTLTOLPDLCPDYIDHFLCE	180
Db	121	DRYAVACRPHLYVIMNWVCFCLRAAASWMLGFGNSVLOSILITLMPRCHOEHEDHFCE	180
Qy	181	VPALILKISCVEHTNEAEFLVSLFPHILPILTLILSYAFIVRVLVIRIOSAEGROKAFGT	240
Db	181	VPALILKISCADTKPIEALFEFFSVAILLIPVLLILISYGLFAQVNLVIRSAEGROKAFGT	240
Qy	241	CGSHLIVLSLEFYSTAVSYLLOPPSPSSKDOCKWVSLFYGIILAPMLNPLIYTLRNKEVKEG	300
Db	241	CGSHMIVVLSLEFYGTAYMYLLOPPSPSTSKDMCKWVSLFYGIITMSLNLIIYSLRNKDMKEA	300
Qy	301	EKRILVAREYELIK 313	
Db	301	EKRILMPRIEFCKK 313	

```

RESULT 11
PCT-US02-03635-48
: Sequence 48. Application PC/TUS0203635
:
: GENERAL INFORMATION:
:
: APPLICANT: INCYTE GENOMICS, INC.
: APPLICANT: BAUGHN, Mariah R.
: APPLICANT: TRIBOLET, Catherine M.
: APPLICANT: NGUYEN, Daniel B.
:
: APPLICANT: THORNTON, Michael
:
: APPLICANT: VAO, Monique G.
:
: APPLICANT: KALLICK, Deborah A.
: APPLICANT: GANDHI, Ameena R.
: APPLICANT: WALIA, Nandier K.
: APPLICANT: ARIZO, Chandra
: APPLICANT: ELLIOTT, Vicki S.
: APPLICANT: HAFALIA, April J.A.
: APPLICANT: RAMKUMAR, Jayalaxmi
: APPLICANT: JIN, Pei
: APPLICANT: TANG, Y.Tom
: APPLICANT:
: YUE, Henry
:

```

```

; APPLICANT: REDDY, Roopa
; APPLICANT: BUREFORD, Neil
; APPLICANT: LU, Dyung Alina M.
; APPLICANT: GRAUL, Richard C.
; APPLICANT: KAHN, Farrah A.
; APPLICANT: WALSH, Roderick T.
; APPLICANT: ISON, Craig H.
; APPLICANT: RICHARDSON, Thomas W.
; APPLICANT: GRIFFIN, Jennifer A.
; APPLICANT: WARREN, Bridget A.
; APPLICANT: YANG, Junming
; APPLICANT: LEE, Ernestine A.
; APPLICANT: HARLAND, Lee
; TITLE OF INVENTION: G-PROTEIN COUPLED RECEPTORS
; FILE REFERENCE: PI-0357 PCT
; CURRENT APPLICATION NUMBER: PCT/US02/03635
; PRIOR APPLICATION NUMBER: 60/267,322; 60/271,215; 60/274,551; 60/278,507;
; 60/280,597; 60/281,107; 60/282,121
; PRIOR FILING DATE: 2001-02-07; 2001-02-23; 2001-03-08; 2001-03-23;
; 2001-03-30; 2001-04-02; 2001-04-06
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE: PERL Program
; SEQ ID NO 48
; LENGTH: 318
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc-feature
; OTHER INFORMATION: Incyte ID No: 55019501CD1
PCT-US02-03635-48

```

```

Query Match      58.9%; Score 947; DB 1; Length 318;
Best Local Similarity 59.7%; Pred. No. 6,7e-68;
Matches 178; Conservative 49; Mismatches 71; Indels 0; Gaps 0;

QY 10 QETILIGSDRPMLEPPLVFLISTVTYIFGNLTIIISRLDTKHTMYFFLTNLSL 69
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 14 KDFILIGISRPWLEPVLVFLYILMLGINSIIIVSQDLDPOLHSPMYIFLSHLSFL 73
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 70 DICTYCTVPMVNCSTRKIVSYRGCVNQLFIPLALGATEVLLAVMSFDMFVAICRP 129
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 74 DICTYTTTVPOMLFNMGSSCKTISYGGCTVQVAFHMLGTECVLAAMALDRVALCEP 133
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 130 LHYSVIMHORLCLOLAASWVTGFSNSVWLSTLTLOPLCDPYVIDHFLCEVPALLKLSL 189
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 134 LRAVILIMHRPLCOGLVAMAMLSGFSNLSVQVILLVQLPFCGRQVANNFCEVPAVAMIKLSL 193
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 190 VETTANEAEFLVSELPHILPTLLIISAFIVRAVLRIQSAEGROKAGTCSHLIVS 249
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 194 ADTTANDATLAVLVAFVFLVPLALLISYGFIRAVWRIOSSRGHRKAFGTCSSHLIVS 253
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 250 LEYSTAVSVILOPPSSSKDQKMSLFGIITAPMLNPLIYTLRNKEVEGKFRIVAR 307
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 254 LFLPALIYMLQPPSSYSQEGKFLISFYIITPLNPFYTLRNKDKGALRLRLAR 311
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

RESULT 12
; US-09-546-986A-2
; Sequence 2, Application US/09546986A
; GENERAL INFORMATION:
; APPLICANT: Powers, Scott
; APPLICANT: Yang, Jianxin
; APPLICANT: Cutler, Gene
; APPLICANT: Tularik Inc.
; TITLE OF INVENTION: Novel G-Protein Coupled Receptors
; FILE REFERENCE: 018781-00472005
; CURRENT APPLICATION NUMBER: US/09/546, 986A
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 09/524,730
; PRIOR FILING DATE: 2000-03-14
; NUMBER OF SEQ ID NOS: 16

```

```

; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 310
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-546-986A-2

```

```

Query Match      58.4%; Score 939; DB 5; Length 310;
Best Local Similarity 58.7%; Pred. No. 2.8e-67;
Matches 175; Conservative 50; Mismatches 73; Indels 0; Gaps 0;

QY 12 FILGSDRPMLEPPLVFLISTVTYIFGNLTIIISRLDTKHTMYFFLTNLSL 71
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 9 FILGSDRPMLEPPLVFLISTVTYIFGNLTIIISRLDTKHTMYFFLTNLSL 68
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 72 CYTTCTVPMVNCSTRKIVSYRGCVNQLFIPLALGATEVLLAVMSFDMFVAICRP 131
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 69 CYTTTTPQVLVNMGSSCKTISYGGCTVQVAFHMLGTECVLAAMALDRIVASCKPLH 128
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 132 YSVIMHORLCLOLAASWVTGFSNSVWLSTLTLOPLCDPYVIDHFLCEVPALLKLSL 191
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 129 YAVIMHRALCOGLVAMAMLSGFSNLSVQVILLVQLPFCGRQVANNFCEVPAVAMIKLSL 188
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 192 TTANEAEFLVSELPHILPTLLIISAFIVRAVLRIQSAEGROKAGTCSHLIVS 251
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 189 TANDTTLAVLVAFVFLVPLALLISYGFIRAVWRIOSSRGHRKAFGTCSSHLIVS 248
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 252 YSTAVSVILOPPSSSKDQKMSLFGIITAPMLNPLIYTLRNKEVEGKFRIVAR 309
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 249 YLPAIYMLQPPSSYSQEGKFLISFYIITPLNPFYTLRNKDKGALRLRLAR 306
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

RESULT 13
PCT-US02-03635-32
; Sequence 32, Application PC/TUS0203635
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: TRIBOULEY, Catherine M.
; APPLICANT: NGUYEN, Dannel B.
; APPLICANT: THORNTON, Michael
; APPLICANT: YAO, Monique G.
; APPLICANT: KALLICK, Deborah A.
; APPLICANT: GANDHI, Ameena R.
; APPLICANT: WALIA, Nandinder K.
; APPLICANT: ARVIZU, Chandra
; APPLICANT: ELLIOTT, Vicki S.
; APPLICANT: HAFALIA, April J.A.
; APPLICANT: RAMKUMAR, Jayalaxmi
; APPLICANT: JIN, Pei
; APPLICANT: TANG, Y. Tom
; APPLICANT: YUE, Henry
; APPLICANT: REDDY, Roopa
; APPLICANT: BUREFORD, Neil
; APPLICANT: LU, Dyung Alina M.
; APPLICANT: GRAUL, Richard C.
; APPLICANT: KAHN, Farrah A.
; APPLICANT: WALSH, Roderick T.
; APPLICANT: ISON, Craig H.
; APPLICANT: RICHARDSON, Thomas W.
; APPLICANT: GRIFFIN, Jennifer A.
; APPLICANT: WARREN, Bridget A.
; APPLICANT: YANG, Junming
; APPLICANT: LEE, Ernestine A.
; APPLICANT: HARLAND, Lee
; TITLE OF INVENTION: G-PROTEIN COUPLED RECEPTORS
; FILE REFERENCE: PI-0357 PCT
; CURRENT APPLICATION NUMBER: PCT/US02/03635
; CURRENT FILING DATE: 2002-02-06
; PRIOR APPLICATION NUMBER: 60/267,322; 60/271,215; 60/274,551; 60/278,507;
; 60/280,597; 60/281,107; 60/282,121
; PRIOR FILING DATE: 2001-02-07; 2001-02-23; 2001-03-08; 2001-03-23;

```

```

; 2001-03-30; 2001-04-02; 2001-04-06
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE: PERL Program
; SEQ ID NO 32
; LENGTH: 309
; TYPE: PRF
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: 7477359CD1
PCT-US02-03635-32

Query Match      58.2%; Score 936; DB 1; Length 309;
Best Local Similarity 59.3%; Pred. No. 4.9e-67;
Matches 179; Conservative 51; Mismatches 72; Indels 0; Gaps 0;

QY 5 NDSIOEFLIGFSRDMLEFPLLYVFLISYTYTIFGNTLIIVSRDLTHHPMFEFL 64
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 5 NDSIMDFILIGFSRDMLEFPLLYVFLISYTYTIFGNTLIIVSRDLTHHPMFEFL 64
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 65 NLSLDCYTTCTVPQMLVNLCSIRKVISYRGCAQLFIFLAGATEYLLAVMSFDMFV 124
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 65 NLSLDCYTTCTSLAPQVLNMLQRPKRTTYGGCAQLYISLAGSTECILLADMDRYI 124
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 125 AICRPLHYVIMHORLCLQAAASWTFGSNSWLSLTITQLPLCDPYVIDHFLCEVPAL 184
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 125 AVCKPLHYVIMPRCLQQAASISWLSGLASSLIHAFTIQLPLCGNHRDLHFICEVPAL 184
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 185 LKLSCEVTANEAELFVLSLFLPLTLISYAFIVRAVLRQSAEGROKAFGTGSH 244
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 185 LKLAQVDTYNELVLFVSVLFPVITPALISISYGFITQAVLRKISVEARHKAFTSCSSH 244
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 245 LTVVSLFYSTAIVSYIOLPPSPSSKDGKMSLFYGIAPMLNPLIYTLRNKKEGFKRL 304
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 245 LTVVLIIFYGTIIYVLIQPSDSYAQDQCKFSLFYTWTPTLNPILYTLRNKDKMEALRKL 304
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 305 VA 306
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 305 LS 306
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 14
US-09-965-422-12
; Sequence 12, Application US/09965422
; GENERAL INFORMATION:
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Casman, Stacie
; APPLICANT: Padigar, Muralidhara
; APPLICANT: Dickson, Kevin
; APPLICANT: Vernet, Corine
; APPLICANT: Spaderina, Steven K
; APPLICANT: Shenoy, Suresh G
; APPLICANT: Gerlach, Valerie
; APPLICANT: Ellerman, Karen
; APPLICANT: Edinger, Shlomit
; APPLICANT: MacDougall, John R
; APPLICANT: Smithson, Glenda
; APPLICANT: Li, Li
; APPLICANT: Malyankar, Urjal M
; APPLICANT: Taylor, Sarah
; APPLICANT: Gunther, Erik
; APPLICANT: Tchernev, Velizar T
; TITLE OF INVENTION: Novel Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21401-132
; CURRENT APPLICATION NUMBER: US/09/965,422
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: 60/236,286
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: 60/236,284
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: 60/237,581
; PRIOR FILING DATE: 2000-10-03
```

```

; PRIOR APPLICATION NUMBER: 60/238,735
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/240,736
; PRIOR FILING DATE: 2000-10-16
; PRIOR APPLICATION NUMBER: 60/260,019
; PRIOR FILING DATE: 2001-01-05
; PRIOR APPLICATION NUMBER: 60/260,338
; PRIOR FILING DATE: 2001-01-08
; PRIOR APPLICATION NUMBER: 60/262,156
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 60/262,498
; PRIOR FILING DATE: 2001-01-18
; PRIOR APPLICATION NUMBER: 60/263,133
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: 60/263,691
; PRIOR FILING DATE: 2001-01-24
; PRIOR APPLICATION NUMBER: 60/266,109
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 60/271,634
; PRIOR FILING DATE: 2001-02-26
; NUMBER OF SEQ ID NOS: 127
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 309
; TYPE: PRF
; ORGANISM: Homo sapiens
US-09-965-422-12

Query Match      58.2%; Score 936; DB 5; Length 309;
Best Local Similarity 59.3%; Pred. No. 4.9e-67;
Matches 179; Conservative 51; Mismatches 72; Indels 0; Gaps 0;

QY 5 NDSIOEFLIGFSRDMLEFPLLYVFLISYTYTIFGNTLIIVSRDLTHHPMFEFL 64
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 5 NDSIMDFILIGFSRDMLEFPLLYVFLISYTYTIFGNTLIIVSRDLTHHPMFEFL 64
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 65 NLSLDCYTTCTVPQMLVNLCSIRKVISYRGCAQLFIFLAGATEYLLAVMSFDMFV 124
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 65 NLSLDCYTTCTSLAPQVLNMLQRPKRTTYGGCAQLYISLAGSTECILLADMDRYI 124
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 125 AICRPLHYVIMHORLCLQAAASWTFGSNSWLSLTITQLPLCDPYVIDHFLCEVPAL 184
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 125 AVCKPLHYVIMPRCLQQAASISWLSGLASSLIHAFTIQLPLCGNHRDLHFICEVPAL 184
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 185 LKLSCEVTANEAELFVLSLFLPLTLISYAFIVRAVLRQSAEGROKAFGTGSH 244
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 185 LKLAQVDTYNELVLFVSVLFPVITPALISISYGFITQAVLRKISVEARHKAFTSCSSH 244
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 245 LTVVSLFYSTAIVSYIOLPPSPSSKDGKMSLFYGIAPMLNPLIYTLRNKKEGFKRL 304
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 245 LTVVLIIFYGTIIYVLIQPSDSYAQDQCKFSLFYTWTPTLNPILYTLRNKDKMEALRKL 304
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 305 VA 306
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 305 LS 306
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 15
US-10-032-189-111
; Sequence 111, Application US/10032189
; GENERAL INFORMATION:
; APPLICANT: Alsobrook II, John P
; APPLICANT: Tchernev, Velizar T
; APPLICANT: Liu, Xiaohong
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Zerhusen, Bryan D
; APPLICANT: Patlurajan, Meera
; APPLICANT: Grosse, William M
; APPLICANT: Lepley, Denise M
; APPLICANT: Burgess, Catherine E
; APPLICANT: Shinkets, Richard A
; APPLICANT: Grosse, William M
```

```

1  APPLICANT: Szekeres, Edward S
2  APPLICANT: Vernet, Corine A.M.
3  APPLICANT: Li, Li
4  APPLICANT: Casman, Stacie J
5  APPLICANT: Boldog, Ferenc L
6  APPLICANT: Gorman, Linda
7  APPLICANT: Gangolli, Esna A
8  APPLICANT: Fernandes, Elma R
9  APPLICANT: Rieger, Daniel K
10 APPLICANT: Edinger, Shlomit R
11 APPLICANT: Gunther, Erik
12 APPLICANT: Millet, Isabelle
13 APPLICANT: Sciore, Paul
14 APPLICANT: Ellerman, Karen
15 APPLICANT: MacDougall, John R
16 APPLICANT: Smithson, Glenda
17 TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
18 FILE REFERENCE: 21402-228
19 CURRENT APPLICATION NUMBER: US/10/032,189
20 CURRENT FILING DATE: 2001-12-21
21 PRIOR APPLICATION NUMBER: 60/257,495
22 PRIOR FILING DATE: 2000-12-21
23 PRIOR APPLICATION NUMBER: 60/258,171
24 PRIOR FILING DATE: 2000-12-20
25 PRIOR APPLICATION NUMBER: 60/269,940
26 PRIOR FILING DATE: 2001-02-20
27 PRIOR APPLICATION NUMBER: 60/274,192
28 PRIOR FILING DATE: 2001-03-08
29 PRIOR APPLICATION NUMBER: 60/277,826
30 PRIOR FILING DATE: 2001-03-22
31 PRIOR APPLICATION NUMBER: 60/279,840
32 PRIOR FILING DATE: 2001-03-29
33 PRIOR APPLICATION NUMBER: 60/282,981
34 PRIOR FILING DATE: 2001-04-11
35 PRIOR APPLICATION NUMBER: 60/283,656
36 PRIOR FILING DATE: 2001-04-13
37 PRIOR APPLICATION NUMBER: 60/309,247
38 PRIOR FILING DATE: 2001-07-31
39 PRIOR APPLICATION NUMBER: 60/311,754
40 PRIOR FILING DATE: 2001-08-17
41 PRIOR APPLICATION NUMBER: 60/313,331
42 PRIOR FILING DATE: 2001-08-17
43 NUMBER OF SEQ ID NOS: 260
44 SOFTWARE: PatentIn Ver. 2.1
45 SEQ ID NO 111
46 LENGTH: 309
47 TYPE: prt
48 ORGANISM: Homo sapiens
49 US-10-032-189-111

```

Query Match	58.2%;	Score 936;	DB 6;	Length 309;
Best Local Similarity	59.3%;	Pred. No. 4.9e-67;		
Matches 179;	Conservative 51;	Mismatches 72;	Indels 0;	Gaps 0;

[illegible]

Db	2.5	LTVAIIIEGTLIIYYLQPSDSYADQCKFISLEFYTMTPTLNPILIIYLRKKDKKALRLK	304
Qy	305	VA 306	
	::		
Db	305	LS 306	

Search completed: August 25, 2002, 20:15:19
Job time: 482 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 25, 2002, 14:56:02 ; Search time 1869.44 Seconds

(without alignments)
10555.955 Million cell updates/sec

Title: US-09-800-321A-3

Perfect score: 943
Sequence: 1 tatgaattgggtaaatgaca.....tcttctaatacaagaataa 943

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl:*
1: gb.ba:*
2: gb.hg:*
3: gb.in:*
4: gb.om:*
5: gb.ov:*
6: gb.pat:*
7: gb.ph:*
8: gb.pl:*
9: gb.pr:*
10: gb.ro:*
11: gb.sts:*
12: gb.sy:*
13: gb.un:*
14: gb.vl:*
15: em.ba:*
16: em.fun:*
17: em.hum:*
18: em.in:*
19: em.mu:*
20: em.om:*
21: em.or:*
22: em.ov:*
23: em.pat:*
24: em.ph:*
25: em.pl:*
26: em.ro:*
27: em.sts:*
28: em.un:*
29: em.vl:*
30: em.hg.hum:*
31: em.hg.inv:*
32: em.hg.other:*
33: em.hggo.inv:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match Length	ID	Description
------------	-------	-------	--------------	----	-------------

1	943	100.0	943	6	AX282353	Sequence
2	940.4	99.7	993	6	AX282351	Sequence
3	940.4	99.7	44788	6	AL133267	Human DNA
4	940.4	99.7	176277	2	AC025941	Sequence
5	938.8	99.6	942	6	AX195281	Sequence
6	938.8	99.6	1488	6	AX195283	Sequence
7	937.4	99.4	939	6	AX241843	Sequence
8	937.2	99.4	956	6	AX282355	Sequence
9	727.8	77.2	1074	6	AX241494	Sequence
10	726.2	77.0	1071	6	AX241494	Sequence
11	726.2	77.0	1071	6	AX242075	Sequence
12	726.2	77.0	1074	9	HS302584	Sequence
13	726.2	77.0	1074	9	HS302585	Sequence
14	726.2	77.0	1074	9	HS302587	Sequence
15	726.2	77.0	1074	9	HS302587	Sequence
16	726.2	77.0	1074	9	HS302588	Sequence
17	726.2	77.0	1074	9	HS302589	Sequence
18	726.2	77.0	1074	9	HS302590	Sequence
19	726.2	77.0	1074	9	HS302591	Sequence
20	726.2	77.0	1074	9	HS302592	Sequence
21	726.2	77.0	100375	9	HS193812	Sequence
22	726.2	77.0	166758	2	AC024428	Sequence
23	726.2	77.0	176277	2	AC025941	Sequence
24	666.8	70.7	942	10	AY073555	Sequence
25	665.2	70.5	185331	2	AL589742	Sequence
26	647.4	68.7	648	9	U86270	Sequence
27	647.4	68.0	648	6	AX242056	Sequence
28	630.6	66.9	648	9	U86271	Sequence
29	610.2	64.7	646	9	U86275	Sequence
30	603.8	64.0	161043	10	AL589651	Sequence
31	602.8	63.9	942	10	AY073554	Sequence
32	600.6	63.7	161043	10	AL589651	Sequence
33	590.4	62.6	940	10	AY074039	Sequence
34	588.4	62.4	1214	10	RATOLIRECE	Sequence
35	535.6	56.8	942	9	HS302546	Sequence
36	534	56.6	942	9	HS302537	Sequence
37	534	56.6	942	9	HS302538	Sequence
38	534	56.6	942	9	HS302539	Sequence
39	534	56.6	942	9	HS302540	Sequence
40	534	56.6	942	9	HS302541	Sequence
41	534	56.6	942	9	HS302542	Sequence
42	534	56.6	942	9	HS302543	Sequence
43	534	56.6	942	9	HS302544	Sequence
44	534	56.6	942	9	HS302545	Sequence
45	534	56.6	134463	9	AL662791	Human DNA

ALIGNMENTS

RESULT 1
AX282353
LOCUS AX282353 943 bp DNA
DEFINITION Sequence 3 from Patent WO0166746.
ACCESSION AX282353
VERSION AX282353.1 GI:16609543
KEYWORDS
SOURCE
ORGANISM human.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS Padigaru,M., Burgess,C.E., Majumder,K., Mishra,V.S., Li,L.,
Baumgartner,J.C., Spytek,K.A. and Tchernev,V.T.
TITLE G-protein coupled receptor related polypeptides
JOURNAL Patent: WO 0166746-A 3 13-SEP-2001,
Curagen Corporation (US)
FEATURES
source
1. 943
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"

BASE COUNT 221 a 227 c 187 g 308 t
ORIGIN

Query Match 100.0%; Score 943; DB 6; Length 943;
Best Local Similarity 100.0%; Pred. No. 2.4e-257;
Matches 943; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 tatgaattggtaataagacagcatcatcagaggtttatctctggtgttcagatcg 60
   |||||||
DB 1 tatgaattgggttaaatgacagcatcatcagaggtttatctctggtgttcagatcg 60
QY 61 acctgagctgaggttccacatctctgtgtctcttgattcttaccctgtgacatctt 120
   |||||||
DB 61 acctgagctgaggttccacatctctgtgtctcttgattcttaccctgtgacatctt 120
QY 121 tggcaactgacatattatctagtgcaagccttgacacacaaactcatatccccatgta 180
   |||||||
DB 121 tggcaactgacatattatctagtgcaagccttgacacacaaactcatatccccatgta 180
QY 181 ttttttcttcaacatcatcatcactcctgtgactctgtttacacacatgtaacgtccaca 240
   |||||||
DB 181 ttttttcttcaacatcatcatcactcctgtgactctgtttacacacatgtaacgtccaca 240
QY 241 aatgctagtaaatattatgacagcatcagaaagtaacatcagtacgtggttgaaccca 300
   |||||||
DB 241 aatgctagtaaatattatgacagcatcagaaagtaacatcagtacgtggttgaaccca 300
QY 301 gctttcataattcttgaccttgagggtactgataatctctcctgtgacgtcgtcgtc 360
   |||||||
DB 301 gctttcataattcttgaccttgagggtactgataatctctcctgtgacgtcgtcgtc 360
QY 361 tgaattggtttagtactattgtgtgacctcacttactcagttatacagacagagact 420
   |||||||
DB 361 tgaattggtttagtactattgtgtgacctcacttactcagttatacagacagagact 420
QY 421 ctgcccacagttgagagcgtgacatcctgaggttactggttttagtaacatcagttggtgtgc 480
   |||||||
DB 421 ctgcccacagttgagagcgtgacatcctgaggttactggttttagtaacatcagttggtgtgc 480
QY 481 taacctgactctcagcgtgacatctgtgacacctatgtatagatcaactctctgtga 540
   |||||||
DB 481 taacctgactctcagcgtgacatctgtgacacctatgtatagatcaactctctgtga 540
QY 541 agtccctgacgtctcaagttatctgtgtgtgagacagacagaaagagagctgaactatt 600
   |||||||
DB 541 agtccctgacgtctcaagttatctgtgtgtgagacagacagaaagagagctgaactatt 600
QY 601 ccttgcaagtgagctcttccatctaatcaccctgacacatcctctatatacatgcttt 660
   |||||||
DB 601 ccttgcaagtgagctcttccatctaatcaccctgacacatcctctatatacatgcttt 660
QY 661 taattgctgagagtaattgagatatacagctgctgaagtcgacaaaaagcatttggagc 720
   |||||||
DB 661 taattgctgagagtaattgagatatacagctgctgaagtcgacaaaaagcatttggagc 720
QY 721 atgtgttccacataattggtgtgtctctttttagtaagcagcgtctctgtgaact 780
   |||||||
DB 721 atgtgttccacataattggtgtgtctctctttttagtaagcagcgtctctgtgaact 780
QY 781 gcaaacacatctgcccagcgtcccaagaaacaaaggaatggttctctcttcatatggaat 840
   |||||||
DB 781 gcaaacacatctgcccagcgtcccaagaaacaaaggaatggttctctcttcatatggaat 840
QY 841 catgtacccaatgctgaatcccttatatacatcacttaggaacaagaggttaagaag 900
   |||||||
DB 841 catgtacccaatgctgaatcccttatatacatcacttaggaacaagaggttaagaag 900
QY 901 ctttaaaaggtgtgtgcaagagctctcttaacaaagaataa 943
   |||||||
DB 901 ctttaaaaggtgtgtgcaagagctctcttaacaaagaataa 943
```

RESULT 2
AX282351

LOCUS AX282351 993 bp DNA linear PAT 02-NOV-2001
DEFINITION Sequence 1 from Patent WO0166746.
ACCESSION AX282351
VERSION AX282351.1 GI:16609542
KEYWORDS
SOURCE
ORGANISM human.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS Padigaru, M., Burgess, C.E., Majumder, K., Mishra, V.S., Li, L.,
Baumgartner, J.C., Spytek, K.A. and Tchernev, Y.T.
TITLE G-protein coupled receptor related polypeptides
JOURNAL Patent: WO 0166746-A 1 13-SEP-2001;
Curagen Corporation (US)
FEATURES
source 1..993
Location/Qualifiers
BASE COUNT 246 a 233 c 194 g 320 t
ORIGIN

Query Match 99.7%; Score 940.4; DB 6; Length 993;
Best Local Similarity 99.9%; Pred. No. 1.3e-256;
Matches 941; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
QY 2 atgaattggtaataagacagcatcatcagaggtttatctctggtgttcagatcga 61
   |||||||
DB 15 atgaattgggttaaatgacagcatcatcagaggtttatctctggtgttcagatcga 74
QY 62 ccttgctgaggttccacatctctgtgtctcttctgattcttaccatgtaacatctt 121
   |||||||
DB 75 ccttgctgaggttccacatctctgtgtctcttctgattcttaccatgtaacatctt 124
QY 122 ggaactctgacatattctagtgtaagcgttgaaacaaacttaaccacatgta 181
   |||||||
DB 135 ggaactctgacatattctagtgtaagcgttgaaacaaacttaaccacatgta 194
QY 182 ttttcttcaacatcatcacccttgagatctctgttacaacacatgtaacgtccaca 241
   |||||||
DB 195 ttttttcttcaacatcatcacccttgagatctctgttacaacacatgtaacgtccaca 254
QY 242 atgtcagtaaatattatgacatcagaaagtaatacagttatctgtgtgtgagccag 301
   |||||||
DB 255 atgtcagtaaatattatgacatcagaaagtaatacagttatctgtgtgtgagccag 314
QY 302 ctcttcaatttctgacctgagggtactgtaatactctctctgacctgacatgcttt 361
   |||||||
DB 315 ctcttcaatttctgacctgagggtactgtaatactctctctgacctgacatgcttt 374
QY 362 gatgtgtttagtaatttctgacctctcacttactcagttatcaatgacacagagact 421
   |||||||
DB 375 gatgtgtttagtaatttctgacctctcacttactcagttatcaatgacacagagact 434
QY 422 tgcctcagttgacatcgtatctcgtgttactcgtgttttagtaacacagtggtgtct 481
   |||||||
DB 435 tgcctcagttgacatcgtatctcgtgttactcgtgttttagtaacacagtggtgtct 494
QY 482 acctgactctcagcgtcgaactctgtgaacccctatgataatatacacttctctgaa 541
   |||||||
DB 495 acctgactctcagcgtcgaactctgtgaacccctatgataatatacacttctctgaa 554
QY 542 gtccctgacatgctcaagttatctgtgtgagacaacagcaaaagagctgaactatc 601
   |||||||
DB 555 gtccctgacatgctcaagttatctgtgtgagacaacagcaaaagagctgaactatc 614
QY 602 ctgtcagtgagctcttccatcctaatacccttgaaacatccttatcatatgcttt 661
   |||||||
DB 615 ctgtcagtgagctcttccatcctaatacccttgaaacatccttatcatatgcttt 674
QY 662 atgtcagagcagttatgagatagacagctgctgaaagtcgacaaagcatttggaca 721
   |||||||
```

Db 675 ATTGTCGACAGATGATGAGATACAGTCTGTCGTAAGTCGACAAAAAGCATTTGGACCA 734

QY 722 TgTgttcccatcctaattgTgtgtctcttttatagtaacgcgctctgtacctg 781
|||||
Db 735 TGTGTTCCCATCTAATTGTTGTCTCTTTTATATAGTACAGCGTCTGTACTCTG 794

QY 782 caacacacttcgcgcagctcccaaggaaggaagatggttctctctctatgtaac 841
|||||
Db 795 CAACACACTTTCGCCACTCCAAAGCAAGAAAGATGTTCTCTCTATGGAATC 854

QY 842 atgacaccatgctgtaatcccttatatactaggaacagagagtaaaaggaagc 901
|||||
Db 855 ATTGCACCCATGCTGATTCCTTATATATACACTTTCGACAAAGAGGTAAGGAGGC 914

QY 902 tttaaaagttgTgtgaagaagctctcttaatacaagaataaa 943
|||||
Db 915 TTTAAAGGTTGTTGCAAGGCTCTTATATCAAGAAATTA 956

RESULT 3
AL133267
LOCUS
DEFINITION
AL133267 44788 bp DNA linear PRI 10-SEP-2001
Human DNA sequence from clone RP3-408B20 on chromosome 6. Contains
a gene and two pseudogenes for 7 transmembrane receptor (rhodopsin
family) (olfactory receptor like) proteins and a gene for a novel
protein similar to 60S acidic ribosomal protein P2 (RPLP2), ESTs,
STS and GSSs, complete sequence.
AL133267
AL133267.9 GI:10185396
HG: 7 transmembrane; olfactory receptor; ribosomal protein; RPLP2.
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 44788)
WILLIAMS,S.
Direct Submission
Submitted (20-AUG-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk
Requests: clonerequest@sanger.ac.uk
On Sep 18, 2000 this sequence version replaced gi:9588470.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information
on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
was generated from part of bacterial clone contigs of human
chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping
Group. Further information can be found at
http://www.sanger.ac.uk/HGP/Chr6
IMPORTANT: This sequence is not the entire insert of clone
RP3-408B20. It may be shorter because we sequence overlapping
sections only once, except for a 100 base overlap.
The true left end of clone RP1-11316 is at 44689 in this sequence.
The true right end of clone RP1-193B12 is at 100 in this sequence.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >= 30).
An attempt was made to resolve all sequencing problems, such
as compressions and repeats: all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest. RP3-408B20 is from
the library RP1-3 constructed by the group of Pieter de Jong. For
further details see
http://www.chori.org/Bacpac/home.htm
VECTOR: pCYPAC2.
Location/Qualifiers

source
1. .44788
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="6"
/clone="RP3-408B20"
/clone_11b="RP1-3"
/complement(1..100)
/note="match: STS: Em:HS193B12T"
222..535
/note="AluXs repeat: matches 3..301 of consensus"
932..1870
/gene="d4408B20.1"
932..1870
/gene="d4408B20.1"
/note="d4408B20.1 (7 transmembrane receptor (rhodopsin
family) (olfactory receptor like) pseudogene (hs6M1-33P))
match: proteins: Tr:07600 Tr:07601 Tr:07602 Tr:035434
Sw:015062 Tr:09WVD9 Sw:013607 Tr:090806 Tr:063394
Tr:09WY13 Tr:095918 Tr:091084 Tr:095006 Tr:091399
Tr:09Q222 Sw:095156 Sw:095157 Tr:0906F6 Tr:091299
Tr:090JL5 Tr:062007 Sw:P23275 Tr:095371 Tr:090JAO
Tr:090218 Tr:090219 Tr:09PSJ3 Tr:P70526 Tr:09Q221"
/codon_start=1
/pseudo
/evidence=not_experimental
complement(1518..1963)
/note="match: GSS: Em:AQ456128"
2024..2298
/note="AluJo repeat: matches 1..280 of consensus"
2299..2382
/note="42 copies 2 mer ag 79% conserved"
2359..2650
/note="2 copies 46 mer 91% conserved"
complement(4227..4576)
/note="match: STS: Em:G31214"
4719..5074
/note="L2 repeat: matches 2375..2750 of consensus"
complement(4834..5686)
/note="match: GSS: Em:AQ748194"
5098..5236
/note="L2 repeat: matches 1381..1525 of consensus"
5707..5843
/note="L2 repeat: matches 102..240 of consensus"
6275..6509
/note="L2 repeat: matches 5093..5329 of consensus"
complement(6590..6889)
/note="match: GSS: Em:U94441"
6679..6752
/note="MER76 repeat: matches 613..685 of consensus"
6753..6850
/note="L1P repeat: matches 1..98 of consensus"
6820..7010
/note="L1P repeat: matches 580..776 of consensus"
7006..7615
/note="L1P repeat: matches 900..1509 of consensus"
7611..9145
/note="L1P repeat: matches 3813..5344 of consensus"
9751..10411
/note="HERV1 repeat: matches 3244..3919 of consensus"
10555..10663
/note="L1R16A repeat: matches 203..307 of consensus"
10964..11131
/note="L1R1H repeat: matches 1..168 of consensus"
11705..12141
/note="L1M4 repeat: matches 3898..4367 of consensus"
13143..13516
/note="match: GSS: Em:AQ100659"
13668..13803
/note="MIR repeat: matches 58..199 of consensus"
complement(15225..15506)
/note="match: GSS: Em:AQ100750"
15627..16228
/note="match: GSS: Em:AQ544878"

```
repeat_region 16824..17089
/note="AluJb repeat: matches 29..311 of consensus"
repeat_region 17090..17158
/note="AluJ/FLAM repeat: matches 1..81 of consensus"
repeat_region 17553..17848
/note="AluSx repeat: matches 3..298 of consensus"
misc_feature 17789..18240
/note="match: STS: Em:G31213"
repeat_region 20021..20249
/note="L1PB1 repeat: matches 5904..6155 of consensus"
20708..21649
/gene="d3408820.2"
20708..21649
/gene="d3408820.2"
/note="match: CDNAS: Em:AF044034 Em:Y14442 Em:X89700
Em:Y07557 Em:AF042023 Em:AF091573 Em:AF091574 Em:X89670
Em:M64391 Em:AF102540 Em:AF010293 Em:M64388 Em:DL2820
Em:AF034896
match: ESTs: Em:N68399
match: proteins: Tr:Q43883 Tr:Q9Q221 Tr:Q921K0 Tr:Q9Q220
Tr:Q63394 Tr:Q95499 Tr:Q9Q219 Sw:P23270 Tr:Q9Y3P7
Tr:Q9Q222 Tr:Q9Y299 Tr:Q62943 Tr:Q90807 Tr:Q90808
Tr:Q9WV13 Tr:Q63395 Tr:Q95918 Sw:P23266 Sw:P37069
Sw:Q95156 Tr:Q9UGF4 Tr:Q9WU86 Tr:Q60403 Tr:Q76000
Sw:P37070 Tr:Q9QW38 Tr:Q77758 Tr:Q9Q217 Tr:Q9R0K3
Tr:Q9R0K4 Tr:Q60412 Tr:Q9Q100 Tr:Q9QW36 Tr:Q9DU40
Sw:P37072 Sw:P37071 Tr:Q9Y3P5"
/codon_start=-1
/evidence=not_experimental
/product="d3408820.2 (novel 7 transmembrane receptor
(rhodopsin family) (olfactory receptor like) protein
(hs6M1-32))"
/protein_id="CAC14158.1"
/db_xref="GI:10944516"
/translation="MMWVNDSTIIOEFLIGESDPMLPEPLLVPLISYNTIFGNLT
ITLVSRDITKHPMYPELNLISLDICVTYCPKMLKSLIRVYSRGVADLP
ITFLAGNTEYLILAVMSFDRFVAICRPLHVSVMHQRCLQLAASWYTFSSWMS
TFLQQLPCDPYIDHFLCEVPALKLSCVETANEELFVSELFHLLPLTILIST
AFIVRAVLRIOSAEGRQKAGTCGSHLIVSLVSTAVSYLOPPSSSKDKMWSL
FVGIILMPLNPLLYTLRNKEKGFRLVAVRLIK"
21073..21162
/note="L1R15 repeat: matches 127..212 of consensus"
complement(21244..21671)
/note="match: GSS: Em:AQ077154"
complement(21438..22000)
/note="match: GSS: Em:A2516219"
21466..21785
/note="match: GSS: Em:A2694245"
complement(21501..22000)
/note="match: GSS: Em:A2516216"
21964..21999
/note="18 copies 2 mer ca 100% conserved"
23289..23944
/note="L1MB2 repeat: matches 5508..6168 of consensus"
24195..24291
/note="FLAM_C repeat: matches 38..136 of consensus"
2448..24781
/note="match: GSS: Em:B75689"
24537..24857
/note="match: GSS: Em:NQ102724"
24739..25051
/note="AluSg repeat: matches 1..310 of consensus"
25407..25944
/note="L1M4 repeat: matches 5206..5763 of consensus"
25945..26050
/note="AluSg/X repeat: matches 195..302 of consensus"
26055..26632
/note="match: GSS: Em:AQ057601"
26069..26126
/note="L1MB7 repeat: matches 6111..6169 of consensus"
26164..26220
/note="L1MB7 repeat: matches 6111..6168 of consensus"
```

```
repeat_region 26221..26391
/note="AluJb repeat: matches 137..306 of consensus"
Query Match 99.7% Score 940.4; DB 9; Length 44788;
Best Local Similarity 99.9%; Pred. No. 1.4e-256;
Matches 941; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2 atgaattggttaattgcaagcatatatacagaagattatctctgctgggtttctcagatcga 61
DB 20708 ATGAAATGGGTAATGACAGCATCATACAGAGATTATTTCTGCTGGGTTCTCAGATCGA 20767
QY 62 ccttgctgtaggttccactccttggtgctcttcttgatcttctacactgtagcatcttt 121
DB 20768 CCTTGGCTGGAGTTTCCACTCCTTGCTGCTCTTGTGATTTCTTACACTGTGACCATCTTT 20827
QY 122 ggcacattgacattatcttagtgcagccttgagaccacaaactcaacccatgat 181
DB 20828 GGCAAATCGACCAATTATTCTAGTGTCAACGCTGACACCAAACTTCATTACCCCATGTAT 20887
QY 182 tttcttaccacatcatcatctcctgagatcttgttaaccacatgtaagtcacaa 241
DB 20888 TTTTCTTACCAATCATATCATCTCTGAGATCTTTGTACACACATGTATACAGTCCACAA 20947
QY 242 atgctagtaattatcagatcaggaagaagtaacagttatcgtgtgctgtgtaccag 301
DB 20948 ATGCTAGTAATTTATCGACATCAGGAAGTAATCATGTTATCTGTGCTGTGTAGCCAG 21007
QY 302 ctttcatattctcgctcttgaggagctactgtaataatctctctcgcgctcatgtcctt 361
DB 21008 CTTTTCATATTTCCTGGGCTTGAGGAGTACGTAATATCTTCCCTGGCGCTCATGCTCTT 21067
QY 362 gattggtttgtagctatttctgagcctctccatctcagatcatcagtcagacagaagctc 421
DB 21068 GATTGTTTGTAGCTATTGTTGCGGCTCTCATACAGTAATCATGACCAAGAGCTC 21127
QY 422 tgcctcagtttgtagagctgcacctcgtggttactggttttagtaactagttgtgtct 481
DB 21128 TGCTTCAAGTTGGCAGCTGCATCTCGGTTTACTGTTTACTAAGTATGTTGTGTCT 21187
QY 482 accctgactctcagctgcagctcctgtaaccctatgtaagataagatacactctctgtgaa 541
DB 21188 ACCCTGACTCTCCAGCTGCACCTCTGACCCCTATGATAGATCATCTTCTCTGAGAA 21247
QY 542 gtccctgactgctccaagtatcttgtgttgtagagaacaggaatgtgctgaactcttc 601
DB 21248 GTCCCTGACTGCTCAAGTTATCTTGTGTGAGACCAACAGCAATGAGGCTGAACCTATTC 21307
QY 602 ctgtcagtagagctcttcacatcatatcccttgacacatcattctatatacattgtcttt 661
DB 21308 CTTGTAGTAGCTCTTCCATCTAATACCTCTGACTCATCCTTATATCATATGCTTTT 21367
QY 662 atgtccgagcagatttgagatacagtcgtgaaagtcgacaaaagaattgggaca 721
DB 21368 ATGTCCGAGCAGATTGAGATACAGTACGTGTAAGGTGCAAAAAGCATTTGGGACA 21427
QY 722 tgtgttccatctaatatgtgtgtctcttttatatgtagcagcgtctcgtgtactgt 781
DB 21428 TGTGTTCCCATCTAATATGTGTGTCTCTTTTATATGTACAGCGCTCTGTGTACTGTG 21487
QY 782 caaccaccttcgcccagctccaagagaccaaagaaatgtttctccttcataagatc 841
DB 21488 CAACCACCTTCGCCACCTCCAAAGACCAAGAAAGATGTTTCTCTTCTATGAGATC 21547
QY 842 attgacacatgtcgaatccctctatataatacattgagaacaaagagagtaagaagc 901
DB 21548 ATTGACACCAATGCGAATCCCTTATATATACATTGGAACAAGAGGATGAAGAGG 21607
QY 902 tttaaaagttgtgtcagaagcttcttcaatcaagaataa 943
DB 21608 TTTAAAGTTGTGTTGCAAGAGTCTCTTATATCAAGAAATAA 21649
RESULT 4
```

AC025941/c
LOCUS AC025941 176277 bp DNA linear HTG 26-MAY-2000
DEFINITION Homo sapiens chromosome 6 clone RP11-635011 map 6, WORKING DRAFT
SEQUENCE, 15 unordered pieces.
AC025941
AC025941.2 GI:7408057
VERSION AC025941.2
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 176277)
Birren, B., Linton, L., Nusbaum, C. and Lander, E.
Homo sapiens chromosome 6, clone RP11-635011
Unpublished
2 (bases 1 to 176277)
Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,
Anderson, S., Baldwin, J., Barna, N., Bastien, V., Bede, F.,
Boguslavsky, L., Bouckhagalter, B., Brown, A., Burkett, C.,
Campopiano, A., Castle, A., Choepel, Y., Colangelo, M., Collins, S.,
Collumore, A., Cooke, P., Dearbellano, K., Dewar, K., Diaz, J.S.,
Dodge, S., Domino, M., Doyle, M., Ferreira, P., Fitzhugh, W., Gage, D.,
Galagan, J., Gardyna, S., Glende, S., Goyette, M., Graham, L.,
Grand-Pierre, N., Grant, G., Hagos, B., Heatford, A., Horton, L.,
Howard, J.C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A.,
Klein, J., Lakoque, K., Lamazares, R., Landers, T., Lehoczy, J.,
Levine, R., Liu, C., Liu, G., Locke, K., MacDonald, P., Marquis, N.,
McCarthy, M., McEwan, P., McGuirk, A., McKernan, K., McSheeters, R.,
McDairlin, J., Menus, L., Mihova, T., Miranda, C., Mlenka, V., Morrow, J.,
Murphy, T., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P.,
O'Neill, D., Oliver, T.M., Oliver, J., Peterson, K., Pierre, N.,
Pisani, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D.,
Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B.,
Strange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,
Testafaye, S., Theodore, J., Turrell, A., Travers, M., Triggillo, J.,
Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J.,
Young, G., Zahoun, J., Zimmer, A. and Zody, M.
Young, G., Zahoun, J., Zimmer, A. and Zody, M.
Direct Submission
Submitted (17-MAR-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Apr 4, 2000 this sequence version replaced g1:7259782.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Project name: L8577
Center project name: 635_O_11
Center clone name: 635_O_11
----- Summary Statistics
Sequencing vector: M13; M77815; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 168929 bases at least Q40
Consensus quality: 172827 bases at least Q30
Consensus quality: 174151 bases at least Q20
Insert size: 176000; agarose-fp
Insert size: 174877; sum-of-contigs
Quality coverage: 4.5 in Q20 bases; agarose-fp
Quality coverage: 4.6 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 15 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence.
* as soon as it is available and the accession number will
* be preserved.
1 2997: contig of 2997 bp in length

2998 3097: gap of 100 bp
3098 5066: contig of 2505 bp in length
5067 5706: gap of 100 bp
5707 9209: contig of 3503 bp in length
9210 9309: gap of 100 bp
9310 13614: contig of 4305 bp in length
13615 13714: gap of 100 bp
13715 20639: contig of 6925 bp in length
20640 20739: gap of 100 bp
20740 26545: contig of 5806 bp in length
26546 26645: gap of 100 bp
26646 37068: contig of 10423 bp in length
37069 37168: gap of 100 bp
37169 48083: contig of 10915 bp in length
48084 48183: gap of 100 bp
48184 59015: contig of 10832 bp in length
59016 59115: gap of 100 bp
59116 70420: contig of 11305 bp in length
70421 70520: gap of 100 bp
70521 80466: contig of 9946 bp in length
80467 80566: gap of 100 bp
80567 92325: contig of 11759 bp in length
92326 92425: gap of 100 bp
92426 110572: contig of 18147 bp in length
110573 110672: gap of 100 bp
110673 131898: contig of 21226 bp in length
131899 131998: gap of 100 bp
131999 176277: contig of 44279 bp in length.
Location/Qualifiers
1. 176277
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="6"
/map="6"
/clone.lib="RP11-635011"
/clone.lib="RP11-635011"
1. 2997
/note="assembly_fragment"
3098. 5066
/note="assembly_fragment"
clone.end:sp6
vector.side:left"
5707. 9209
/note="assembly_fragment"
9310. 13614
/note="assembly_fragment"
13715. 20639
/note="assembly_fragment"
20740. 26545
/note="assembly_fragment"
clone.end:t7
vector.side:right"
26646. 37068
/note="assembly_fragment"
37169. 48083
/note="assembly_fragment"
48184. 59015
/note="assembly_fragment"
59116. 70420
/note="assembly_fragment"
70521. 80466
/note="assembly_fragment"
80567. 92325
/note="assembly_fragment"
92426. 110572
/note="assembly_fragment"
110673. 131898
/note="assembly_fragment"
131999. 176277
/note="assembly_fragment"
BASE COUNT 54493 a 33357 c 33254 g 53772 t 1401 others
ORIGIN

Query Match 99.7%; Score 940.4; DB 2; Length 176277;
Best Local Similarity 99.9%; Pred. No. 1.4e-256;
Matches 941; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 atgaattgggtaaatgacagacatcatcagagattattcttctgctgggtcttcagatcga 61
DB 154174 ATGAAATGGGTAAATGACACATCATACAGAGATTATTTGCTGGGTTTCTCAGATCGA 154115

QY 62 ccttgctggaggttcccaacctgtgtgtcttctgattcttcttaccactgtgaccatctt 121
DB 154114 CTTGGCTGGAGTTTCCACCTCTTGCTGTTCTTCTGATTTCTTACACTGTGACCATTTT 154055

QY 122 ggaacatcgaacatattctagtgtcaagcctgtgaaaccaactcaacccccatgtat 181
DB 154054 GGCNAATCTGCACATTTATTTAGTGTACAGCGCTGGACACCAAACTTCATACCCCATGTAT 153995

QY 182 tttttcttccaatcatcatctactccgtgattcttgttaccaccaatgtataagtcacca 241
DB 153994 TTTTCTTACCAATCATCATCTCTGGATCTTGTATACCAACATGTACAGTCCACAA 153935

QY 242 atgctagtaaatattatgcacacatcagaagaagtaatcagttatctgtgctgtgacccag 301
DB 153934 ATGCTAGTAATATTATGACACATCAGAAAGTATCATGTTATCTGTGCTGTGATGCCAG 153875

QY 302 ctlttcataattcttgccttggggtggtggtactgaatatctctctggtgcgtcatgtcctt 361
DB 153874 CTTTCATATTTCTGGCCTTGGGGGACTAGTAATCTTCTCTGGCCGTCATGCTCTT 153815

QY 362 gatgtgttgaattatcttctggtccttccattaccatcagattatcatgaccacagaagctc 421
DB 153814 GATAGGTTGTACTATTTCTGCGCTCTCCTCATTTACATTAATCATGACACGAGAGCTC 153755

QY 422 tgcctcagtttgcagctgacatcctgtgtactgtgtttagtaactcagtggtgtctc 481
DB 153754 TGCCCTCAGTTGGCAGGTGATCTCTGGGTACAGTTTGTAGTACAGTGTGTGTCT 153695

QY 482 accctgaactcctcagctgtcacctctgtgacctatgtgatatgaatccttccctgtgaa 541
DB 153694 ACCCTGACTCTCCAGCTGTCCACCTGTGACCCCTATGTGATGATCATCTTCTGTGAA 153635

QY 542 gtccctcgaactgtctcaagtatctgtgtgagaacaacaagaatgaggtgaactatc 601
DB 153634 GTCCCTGCACTGCTCAAGTTATTTGTGTGAGACACAAATGAGCTGACATATTC 153575

QY 602 ctgttcagtgagctctctccatctcaataaccctgtgacatcctctatcatatatgtctt 661
DB 153574 CTTGTGAGTGAAGCTCTCCATTAATACCCCTGACACTCATCTTATATCATATGCTTT 153515

QY 662 attgtccgagcagtatgtgagatacagctctgtcgaaggtcgcacaaaagcatttgggaca 721
DB 153514 ATTGTCCGAGCATATTGAGGATACAGTGTGGAAGGTGACAAAAGCATTTGGGACA 153455

QY 722 tgtgtgtcccaatcaattgtgtgtgtctcttctttagatgaacgctctctgtgtacctg 781
DB 153454 TGTGTGTCCCATTTAATTTGTGTGTGTCTCTTTTATATGATGACCGCTCTGTGTGACTG 153395

QY 782 caaccacactcgcgcagctcccaagaaccaagaagaatgttctctctctctctatgtgaatc 841
DB 153394 CAACGACCTTGGCCAGCTCCAGTCCAAAGAACAAAGATGTTCTCTCTCTATGGAATC 153335

QY 842 attgcaaccaatgtcgaatccctctatatacacttagaagaagaaggtlaaagaagc 901
DB 153334 ATTGCAACCATGCTGAATCCCTTATATATACACTTAGGAACAAGAGGTAAGGAAGGC 153275

QY 902 ttttaaaaggttggltgcaagagctctcttaataataaataaa 943
DB 153274 TTTAAAGGTTGGTGCAGAGTCTTCTTAATCAAGAAATAA 153233

RESULT 5
AX195281 AX195281 942 bp DNA linear PAT 28-AUG-2001
LOCUS

DEFINITION Sequence 1 from Patent WO0151634.
ACCESSION AX195281
VERSION AX195281.1 GI:15385832
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 942)
AUTHORS Walke,D.W., Wilganowski,N.L., Turner,C.A., Friedrich,G., Abuln,A.,
Zambrowicz,B. and Sands,A.T.
TITLE Human olfactory receptor and polynucleotides encoding the same
JOURNAL Patent: WO 0151634-A 1 19-JUL-2001;
Lexicon Genetics Incorporated (US)
location/Qualifiers
FEATURES
source
BASE COUNT 222 a 228 c 187 g 305 t
ORIGIN

Query Match 99.6%; Score 938.8; DB 6; Length 942;
Best Local Similarity 99.8%; Pred. No. 3.7e-256;
Matches 940; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 atgaattgggtaaatgacagacatcatcagagattattctgctgggtcttcagatcga 61
DB 1 ATGAATGGGTAAATGACACATCATACAGAGATTATTTGCTGGGTTTCTCAGATCGA 60

QY 62 ccttgctggaggttcccaacctgtgtgtcttctgattcttcttaccactgtgaccatctt 121
DB 61 CTTGGCTGGAGTTTCCACCTCTTGCTGTTCTTCTGATTTCTTACACTGTGACCATCTT 120

QY 122 ggaacatcgaacatattctagtgtcagcctgtgaaaccaactcaacccccatgtat 181
DB 121 GGCNAATCTGCACATTTATTTAGTGTACAGCGCTGTGACACCAAACTTCATACCCCATGTAT 180

QY 182 tttttcttccaatcatcatctactccgtgattcttgttaccacaatgtacagttccacaa 241
DB 181 TTTTCTTACCAATCATCATCTCTGGATCTTGTATACCAACATGTACAGTCCACAA 240

QY 242 atgctagtaaatattatgcacacatcagaagaagtaatcagttatctgtgctgtgacccag 301
DB 241 ATGCTAGTAATATTATGACACATCAGAAAGTATCATGTTATCTGTGCTGTGAGCCAG 300

QY 302 ctlttcataattcttgccttggggtggtggtactgaatatctctctggtgcgtcatgtcctt 361
DB 301 CTTTCATATTTCTGGCCTTGGGGGCTACTGAATATCTTCTCTGGCCGTCATGCTCTT 360

QY 362 gatgtgttgaattatcttgcgcctctccattactcagttatcatgacacagaagctc 421
DB 361 GATAGGTTGTGACATTTGTGCGGCTCTCCATTACATCAATGACACGAGACTC 420

QY 422 tgcctcagtttgcagctgtacatcctgtgtgtactgtgtttagtaacccagtggtgtct 481
DB 421 TGCCCTCAGTTGGCAGCCGATCTGTGGTATCAGGTTTATGTAATCATCAGTGTGTGTC 480

QY 482 accctgaactcctcagctgtcacctctgtgacctatgtatgaatacctctctgtgaa 541
DB 481 ACCCTGACTCTCCAGCTGCACTCTGTGACCCCTATGTGTAATCATCTTCTGTGAA 540

QY 542 gtccctgcactgtctcaagtatctgtgtgagaacaagaacaaatgaggtgaactatc 601
DB 541 GTCCCTGCACTGCTCAAGTTATCTTGTGTGAGAACAAAGCAATGAGGCTGAATATTC 600

QY 602 ctgttcagtgagctcttccatcaataaccctgtgaacatcctctatcatatgtctt 661
DB 601 CTTGTGAGTGAAGCTCTTCCATTAATACCCCTGACACTCATCTTATATCATATGCTTTT 660

QY 662 attgtccgagcagtatgtgagatacagctctgtcgaaggtcgcacaaaagcatttgggaca 721
DB 661 ATTGTCCGAGCATATTGAGGATACAGTCTGTGGAAGGTGACAAAAGCATTTGGGACA 720

QY 722 tctgttcccatctaatgtgtgtctctctttatagtagaccgctctctgtactg 781
|||||
Db 721 TGTGTTCCCATCTTAATGTGTGTCTCTTTTATAGTACAGCGTCTGTGTACTG 780
782 caaccactctgcccagctccaagaagaagaatggttctctctcttactgaatc 841
|||||
Db 781 CAACCACTCTGCCAGCTCCAGAGCAAGAAAGATGTTCTCTCTCTATGGAATC 840
QY 842 atgcaccatgctgaatccctctatatatactactgaagaagaagtaagaagc 901
|||||
Db 841 ATGCACCACTGCTGAATCCCTTATATATACCTTAGAGCAAGAGTAAAGAGCG 900
QY 902 tttaaaaggtgtgttgcgaagctctcttaatacaagaataa 943
|||||
Db 901 TTTAAAGGTTGTGTGCAAGAGCTCTCTTATCAAGAAATAA 942
RESULT 6
AX195283 1488 bp DNA linear PAT 28-AUG-2001
LOCUS AX195283 Sequence 3 from Patent WO0151634.
DEFINITION AX195283
ACCESSION AX195283
VERSION AX195283.1 GI:15385833
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1488)
AUTHORS Walke,D.W., Wilganowski,N.L., Turner,C.A., Friedrich,G., Abulin,A.,
Zambowicz,B. and Sands,A.T.
TITLE Human olfactory receptor and polynucleotides encoding the same
JOURNAL Patent: WO 0151634-A 3 19-JUL-2001;
Lexicon Genetics Incorporated (US)
FEATURES
Source 1. 1488
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 421 a 333 c 254 g 477 t 3 others
ORIGIN
Query Match 99.6%; Score 938.8; DB 6; Length 1488;
Best Local Similarity 99.8%; Pred. No. 3.8e-256;
Matches 940; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 2 atgaattggtaaatgacagcatcatagaggattatctgctggttctcagatcga 61
|||||
Db 198 ATGAATGGTAATGACAGCATCATACAGAGATTATTCGTGGGTTTCTCAGATCGA 257
QY 62 ccttgctggaattccaactctgtgtgtctcttgattcttaactgtgacatctt 121
|||||
Db 258 CCTTGCGTGGAGTTCCACTCTGTGTGTCTCTGTGATTTCTTACAGTCACTCTTT 317
QY 122 ggcacatcgaccattatctagtgtcagcctggaacaaactcacaacccatgat 181
|||||
Db 318 GGCATCTGACCATTAATCTAGTGTGACGCGCTGACACCAACTTCATACCCCATGAT 377
QY 182 ttcttctacaactcatcactcctgtgactctgttcaacacatgacagtcacca 241
|||||
Db 378 TTTTCTTACCAATCTATCACTCTGTGATCTTTGTACACCAATGACAGTCCACAA 437
QY 242 atgctagtaaatatgacagcatcagaagaatcaagtaactcgtgctgtgtagccag 301
|||||
Db 438 ATGCTAGTAATTAATGACAGCATCAGAAAGTAATCACTTATCGGCTGTGAGCCAG 497
QY 302 ctcttcacattctggccttgaggctactgaataatcttccctggccgtatgctctt 361
|||||
Db 498 CTTTTCATATTTCTGCGCTTGAGGCTACTGAAATATCTTCTCTGCGCTCATGTCTTT 557
QY 362 gattgttgttagctattgtcgcctctcatattacatcatcacaagaagctc 421
|||||

Db 558 GATAGTTTGTAGCTATTGTGCGGCTCTCCATTACTAGTATCATGCACAGAGACTC 617
QY 422 tgcctcagattgagcagctgcacactcctggttactggttttagtaactcagtggtgtct 481
|||||
Db 618 TGCCCTCAGTTGGAGCGCGCATCTCGGGTTACTGGTTTATGAATCACTAGTGAGTGTCT 677
QY 482 accctgaactctcagctgcacactctgtgaacccatgtgataagataacttctctgtaa 541
|||||
Db 678 ACCCTGACTCTCCAGCTCCACTCTGTGACCCCATATGTAATGATCACTTCTCTGTGAA 737
QY 542 gtccctgacagctcaaatctctgtgtgttagaacaagaagaatgagcctgaactatc 601
|||||
Db 738 GTCCCTGCACTGCTCAATTAATCTTTGTGTGAGCAACAGCAATGAGCTGAATATTC 797
QY 602 ctgtcagtgagctcttccatcataataaccctgacacatcactatatactatgctt 661
|||||
Db 798 CTGTGCACTGAGCTCTTCATCTAATACCCGACATCATCCCTTAATCATATGCTCTTT 857
QY 662 attgtcgaagcaatattgaagatatacagctctgtgaagtgacaacaaagattggagca 721
|||||
Db 858 ATTGTCCAGCAGATTAATGAGATACAGTCTGCTGAAGCTCGACAAAAGCATTTGGGACA 917
QY 722 tctgttcccatctaatgtgtgtctctctttttagtagaccgctctctgtactg 781
|||||
Db 918 TGTGTTCCCATCTTAATGTGTGTCTCTTTTATATACAGCCGTCTGTGTACTCG 977
QY 782 caaccaccttgcgcccagctccaagaagaagaatgattctctctctctatgaaatc 841
|||||
Db 978 CAACCACTCTGCCAGCTCCAGAGCAAGAAAGATGTTCTCTCTGTATGAAATTC 1037
QY 842 attgcacacatgctgaatccctctatatatactactgaagaagaagtgtaagaagc 901
|||||
Db 1038 ATTCGACCACTGCTGAATCCCTTATATATACCTTAGAACCAAGAGGTAAAGAGGC 1097
QY 902 tttaaaaggtgtgttgcgaagctctcttaatacaagaataa 943
|||||
Db 1098 TTTAAAGGTTGTGTGCAAGAGCTCTTCTTATCAAGAAATAA 1139
RESULT 7
AX241843 939 bp DNA linear PAT 26-SEP-2001
LOCUS AX241843 Sequence 591 from Patent WO0127158.
DEFINITION AX241843
ACCESSION AX241843
VERSION AX241843.1 GI:15798718
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
REFERENCE 1 (bases 1 to 939)
AUTHORS Bellenson,J., Smith,D., Lancet,D., Glusman,G., Fuchs,T. and
Yanai,I.
TITLE Olfactory receptor sequences
JOURNAL Patent: WO 0127158-A 591 19-APR-2001;
Disinfects (US); YEDA RESEARCH AND DEVELOPMENT COMPANY, LTD. (IL)
FEATURES
Source 1. 939
Location/Qualifiers
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="(H38g440 nucleotide)"
BASE COUNT 220 a 227 c 187 g 305 t
ORIGIN
Query Match 99.4%; Score 937.4; DB 6; Length 939;
Best Local Similarity 99.9%; Pred. No. 9.4e-256;
Matches 938; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2 atgaattggtaaatgacagcatcatagaggattatctgctggttctcagatcga 61
|||||
Db 1 ATGAATGGTAATGACAGCATCATACAGAGATTATTCGTGGGTTTCTCAGATCGA 60
QY 62 ccttgctggaattccaactctgtgtgtctcttgattcttaactgtgacatctt 121
|||||

|||||
Db 61 CCGTGGCGAGTTCCACCTCTGTGTCCTTGTGATTTCTACACGTGACCATCTTT 120
Qy 122 ggcacatcgacatattatctagtcagcgcctggacacacacatcatatccccatgat 181
Db 121 GGCATATCGACCATATTATCTAGTGTACAGCTGGACACCAACTTCATACCCCATGTAT 180
Qy 182 ttcttcttaaccaatcatcatcctctgacatcttcttaccacacatgtacagtcaccaca 241
Db 181 TTTTCTTACCATCTATCATCTCTGGATCTTGTATACACCATATGATGCCACAA 240
Qy 242 atgtatgataaattatgcagatcaagaaagaaatcaagttatctgtgctgtgtacccag 301
Db 241 ATGCTAGTAAATTTATGACATCAGAAAGTAATCACTTATCCGTGCTGTGACCCAG 300
Qy 302 ctttcatattctgcgccttgggggactactaatatctctcccgccgcatatctctt 361
Db 301 CTTTTCATATTTCGGCCTTGGGGGCTACATCATCTCTCCCTGGCCGTATGTCTTT 360
Qy 362 gattggtttagctatttgcgcctctccatctactactactatcatgacacagagactc 421
Db 361 GATAGGTTTGTAGCTATTGTGGCCCTCCATTAATCATGATATCATGACACAGACTTC 420
Qy 422 tgcctcaagttggcagctgcacacctgtggttactgttttagtaactcagtggttct 481
Db 421 TGCTCCAGTTGGCAGCTGCATCTCGGGTTACTGGTTTAACTCATAGTGTGTCT 480
Qy 482 accctgacctccagctgcgcacctgtgacacctatgtatagatactatctctgtgaa 541
Db 481 ACCCTGACTCTCCAGCTGCCACTCTGTGACCCCTATGTGATAGATACACTTCTCTGTGAA 540
Qy 542 gtccctgacctgtcagaattatcttctgtgtgagacaacagaacaaatgagctgaactatc 601
Db 541 GTCCCTGACACTGCTCAAGTTATCTTGTGTGAGACAACAGCAAAATGAGCGCTGAACTATTC 600
Qy 602 ctgtcagtgagctcttccatctaatcccttgacactcatccttatactataagctttt 661
Db 601 CTTTTCAGTAGCTCTTCATCTATATACCCCTGACACTCATCTATATCATGCTTTT 660
Qy 662 atgtccgagcagatltgagagatacagctctgtaaggtgcagacaaaagacattgggagca 721
Db 661 ATTTCGAGAGATGTAGAGATACAGCTCTGTAAGTGCACAAAAGCAATTTGGGACA 720
Qy 722 tgtgttccacataatgtgtgtctcttcttatagtacagcgcgtctgtgtacctg 781
Db 721 TGTGTTCCCATCTAATATGTGTGTCTCTTTTATAGTACAGCCGCTCTGTGTACTCTG 780
Qy 782 caaccacctgcgcacagctccaaggaacagaagaagtgttctctcttctatagaaac 841
Db 781 CAACCACTCTGCCCCAGCTCCAAAGACCAAGAAAGATGGTTCTCTCTATGGAATC 840
Qy 842 atgtcacacatgtcgaatccctctatatatacacttaggaacagaaggtataagaaagcc 901
Db 841 ATTTCACACCATGTGATGCCCTTATATATATACCTTAGGAACAAGAGTAAAGAAAGCC 900
Qy 902 tttaaaaggtgtgtcgaagagcttctcttaatacaagaaa 940
Db 901 TTTTAAAGGTTGTGCAAGAGTCTCTTATATCAAGAAA 939

RESULT 8
AX282355 956 bp DNA linear PAT 02-NOV-2001
LOCUS AX282355
DEFINITION Sequence 5 from Patent WO016746.
ACCESSION AX282355
VERSION AX282355.1 GI:16609544
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniala; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (patent)
AUTHORS Padigaru,M., Burgess,C.E., Majumder,K., Mishra,V.S., Li,L.,

|||||
Qy 2 atgaattgggaaatgaaacatataaagaaattatctgcctggggttctcagatcga 61
Db 7 ATGAATTGGTAAATGACAGCATATACAGAGTTATCTGTGGGTTCTCAGATCGA 66
Qy 62 ccttgctgaggttccactcctctgtgtcttcttgatcttcttaccactgtgacacatctt 121
Db 67 CCGTGGCGAGTTTCCACTCTCTGTGTGTCCTTGTGATTTCTTACACTGTACCATCTTT 126
Qy 122 ggcacatcgacatattatctagtcagcgcctggacacacacatcatatccccatgat 181
Db 127 GGCATATCGACCATATTATCTAGTGTACAGCTGGACACCAACTTCATACCCCATGTAT 186
Qy 182 ttcttcttaaccaatcatcatcctctgacatcttctttagtaacacatgtagactccaca 241
Db 187 TTTTCTTACCAATCTATCATCTCTGGATCTTGTGTACACCAATGTACAGTCCACAA 246
Qy 242 atgtatgataaattatgcagatcaggaagaatcaagttatctgtgctgtgtgagccag 301
Db 247 ATGCTAGTAAATTTATGAGAGATCAGGAAGTAATACATGTTATCCGTGCTGTGACCCAG 306
Qy 302 ctttcatattctgcgccttgggggactcgtgaatatctctctgcgcgtatgctctt 361
Db 307 CTTTTCATATTTCGGCCTTGGGGCTACTGAATATCTTCTCTGCGCGTATGTCTTT 366
Qy 362 gattggtttagctatttgcgcctctccatctactactactatcatgacacagagactc 421
Db 367 GATAGGTTTGTAGCTATTGTGGCCTCTCCATTAATCATGATATCATGACACAGACTC 426
Qy 422 tgcctcaagttggcagctgcacacctgtggttactgtgtttagtaactcagtggttct 481
Db 427 TGCTCCAGTTGGCAGCTGCATCTCGGGTTACTGGTTTAACTAGTGTGTGCT 486
Qy 482 accctgacctccagctgcgcacctgtgacacctatgtatagatacacttctctgtgaa 541
Db 487 ACCCTGACTCTCCAGCTGCCACTCTGTGACCCCTATGTGATAGATACACTTCTGTGAA 546
Qy 542 gtccctgacctgcgcacagttatctgtgtgtgagacaacagaacaaatgagcctgaactatc 601
Db 547 GTCCCTGCACTGCTCAAGTTATCTTGTGTGTGAGCAACAGCAAAATGAGCCTGAATTC 606
Qy 602 ctgtcagtgagctcttccatctaatcccttgacactcatccttatactataatgcttt 661
Db 607 CTTGTCAGTAGCTCTTCCATCTAATACCCCTGACACTCATCTTATATCATGCTTTT 666
Qy 722 tgtgttccacataatgtgtgtctcttcttatagtacagcgcgtctgtgtacctg 781
Db 727 TGTGTTCCCATCTAATATGTGTGTCTCTTTTATAGTACAGCCGCTCTGTGTACTCTG 786
Qy 782 caaccacctgcgcacagctccaaggaacagaagaagtgttctctcttcttaagaaac 841
Db 787 CAACCACTCTGCCCCAGCTCCAAAGACCAAGAAAGATGGTTCTCTCTTATGGAATC 846
Qy 842 atgtcacacatgtcgaatccctctatatatacacttaggaacagaaggtataagaaagcc 901

Query Match 99.4%; Score 937.2; DB 6; Length 956;
Best Local Similarly 99.7%; Pred. No. 1,1e-255;
Matches 939; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
ORIGIN
BASE COUNT 227 a 228 c 190 g 311 t
TITLE Baumgartner,J.C., Spytek,K.A. and Tchernev,V.T.
JOURNAL G-protein coupled receptor related polypeptides
Patent: WO 016746-A 5 13-SEP-2001;
FEATURES
source Curagen Corporation (US)
location/Qualifiers
1..956
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 227 a 228 c 190 g 311 t

|||||
Db 847 ATGGACCCAGTGGAACTTATATATACCTTAGGAACAGGAGTAAGAGAGGC 906
Oy 902 tttaaaaggttggtgcaagagctctcttaatacaagaataa 943
|||||
Db 907 TTTAAAGGTTGGTGGCAAGCTCTTCTATATACAGAAATTA 948
RESULT 9
HSA302593 1074 bp DNA linear PRI 06-JAN-2001
LOCUS Homo sapiens 6M1-10*02 gene for olfactory receptor, cell line
DEFINITION KR3598
ACCESSION AJ302593 GI:12054410
VERSION 6M1-10*02 gene; olfactory receptor.
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1074)
AUTHORS Ziegler, A., Ehlers, A., Forbes, S.A., Trowsdale, J.,
Uchanska-Ziegler, B., Volz, A., Younger, R. and Beck, S.
TITLE Polymorphic olfactory receptor genes and HLA loci constitute
extended haplotypes
(In) Kasahara, M. (Ed.):
MAJOR HISTOCOMPATIBILITY COMPLEX-EVOLUTION, STRUCTURE, AND
FUNCTION: 110-130;
Springer-Verlag, Tokyo, Japan (2000)
REFERENCE 2 (bases 1 to 1074)
AUTHORS Ziegler, A., Ehlers, A., Forbes, S.A., Trowsdale, J., Volz, A.,
Younger, R.M. and Beck, S.
TITLE Polymorphisms in olfactory receptor genes: a cautionary note
Unpublished
3 (bases 1 to 1074)
AUTHORS Younger, R.M., Amadou, C., Bethel, G., Ehlers, A., Fischer Lindahl, K.,
Forbes, S.A., Horton, R., Milne, S., Mungall, A.J., Trowsdale, J.,
Volz, A., Ziegler, A. and Beck, S.
TITLE Characterisation of clustered MHC-linked Olfactory Receptor Genes
in Human and Mouse
Unpublished
4 (bases 1 to 1074)
AUTHORS Ehlers, A., Beck, S., Forbes, S.A., Trowsdale, J., Volz, A., Younger, R.
TITLE MHC-linked olfactory receptor loci exhibit polymorphism and
contribute to extended HLA/OK-haplotypes
Unpublished
5 (bases 1 to 1074)
AUTHORS Ehlers, A.
TITLE Direct Submission
Submitted (06-NOV-2000) Ehlers A., Institut fuer Immunogenetik,
Universitaetsklinikum Charite, Spandauer Damm 130, 14050 Berlin,
GERMANY
FEATURES
source location/Qualifiers
1..1074
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="6"
/cell_line="KR3598"
1..1074
/gene="6M1-10*02"
1..1074
/gene="6M1-10*02"
/product="olfactory receptor"
/protein_id="CAC20513.1"
/db_xref="GI:12054411"
/translation="NMNWKSVQEPILLVFSQPMLEIRPFVWFLFSYILTFGNIL
ILVSHDEKLHTPMTFUSNLSLDDCTTSTVPQMLVNICTRKVISGCGVADLE
IFLALSTEDCLLAVMCFDFRVAICRPLHSLIMHRCQLQAASVIGSFSNVLOS
TWTKMPLGCHKHDFCEVPAALIKISCDVTANAELEFFIVLELLIVTLLISY
AFIVQAVLRIOSEGRRAKAGTCGSHLIIVSLFYGTATISMYLQPPSSSDRKMVSL
FCGIAPMLNPLIYTLRNKEVKAERLVAKSLNLEIRNMQISFAKDVIVLYTINE

BASE COUNT 277 a 248 c 203 g 346 t
ORIGIN
Query Match 77.2%; Score 727.8; DB 9; Length 1074;
Best Local Similarity 86.4%; Pred. No. 4,1e-196;
Matches 815; Conservative 0; Mismatches 127; Indels 1; Gaps 1;
Oy 2 atgaattggtlaaatagcagcatcatcagagagttcttcgctgggttctcagatcga 61
|||||
Db 1 ATGAATTTGGGTAATATAGAGTGTCCACAGAGGTTCAFTTCGTAGTTTCTCGATCAA 60
Oy 62 ccttgctgaggttccacatcccttggtgcttcttgatcttcaactgagacatctt 121
|||||
Db 61 CCATGGCTAGAGATTCCACCCCTTGTGATGTTCTGTTTCCATTCATCTGACAACTTT 120
Oy 122 ggaacatgacacattatcagtcagtcagctggaacacaaactctaaccccatgat 181
|||||
Db 121 GGCATCTGACAAATTAATTTCTGTCAAGTGGATTTCAAATCCACACCCCTATGTAC 180
Oy 182 tttttctaccaatcatcatcctctgagatcttctgttacaccacatgtacagtcacca 241
|||||
Db 181 TTTTCTTTGACATCTCTCATCTCTGACCTTGTGTAACCAAGTACAGTTCCACAA 240
Oy 242 atgctagtaaatattatgcagcatcaggaagtaactcagttacgtgctgtgaccag 301
|||||
Db 241 ATGCTGGTAAACATATATGCAACACACAGAAAGTAATGATGCTGTGCGCCAG 300
Oy 302 ctttcatattctggcctgggggctcagtaaatatcttcctggcgctcatgtcctt 361
|||||
Db 301 CTTTTCATTTTCTCGGCTTG6GTTCCACAGAAATGTTCTTCTCGGCGCTCATGTGCTTT 360
Oy 362 gattgttctgagctatttctgagcctctccattactcagttacatcagtcacagaagctc 421
|||||
Db 361 GATAGTGTGAGTATTTGCGGCTCTCCATTCATTCATTAATTCACAGACAGAGCTC 420
Oy 422 tgcctcaggttgagagctgcagctcctggttactggttttagtaactcagtggtgtct 481
|||||
Db 421 TGCCTTCAGTGGGAGCTGCATCTGATTAAGTGGCTTACCAATTCAGTATTCACATCC 480
Oy 482 accctgactctcagcctgcagcactctgagcccatctgtatagatccttctctgaa 541
|||||
Db 481 ACCTGGACACTTAAGATGCCACTGTGTGTCACAAAGAGTCACTTCTTGTGAA 540
Oy 542 gtccctgactgctcagattactctgtgtgagacaacagcaatgagctgaaactatc 601
|||||
Db 541 GTCCCTGCTGTGCAAGTGTGCTGTGTGACACACAGCAATGAGGCTGAACTATTC 600
Oy 602 ctgtcagtgagctctccatctataaccctgacatccatcttatatacatgcttt 661
|||||
Db 601 TTCATCAGTGTGCTATTCCTTAAATACCCGTGACACTCATCTTAATATCGTATGCTTTT 660
Oy 662 atgtcgaagcagatcagagatcagctctgtaaggtgcacaaaagcattgaggaca 721
|||||
Db 661 ATTGTCCAAGCAGTGTGAGAAATCCAGTCTGTCAAGGTGCACAAAGCGATTTGGGACA 720
Oy 722 tgtgttcccatctaatgtgtgtctctttttatagtagcagcgctctgtgtactgt 781
|||||
Db 721 TGTGCTCCCATTAATGTGTGTGCTACTTTTATGTGACAGTATCTCATCTGATACCTG 780
Oy 782 caacacactcgccacagctcacaagacaaagaaagatgttctctctctatcgtgaatc 841
|||||
Db 781 CAACACACTTCACCCAGCTCCAAAGACCGGGGAAAGATGTTTCTCTTGTGGGATTC 840
Oy 842 atttcacacacagctgaaatcccttatatatacacttggaaacaagggatgaagaagggc 901
|||||
Db 841 ATTGACACCATGCTGAATCCCTTATATATACATTAGGAACAAAGAGTAAAGGAAGCC 900
Oy 902 tttaaaaggttggttgc-aagagcttcttaatacaagaataa 943
|||||
Db 901 TTTAAAGGTTGGTGGCAAGAGCTCTTCTTATATCAAGAAATTA 943

```

RESULT 10
AX241494 1071 bp DNA Linear PAT 26-SEP-2001
LOCUS Sequence 242 from Patent WO0127158.
DEFINITION AX241494
ACCESSION AX241494
VERSION AX241494.1 GI:15798369
KEYWORDS
SOURCE
ORGANISM synthetic construct.
REFERENCE synthetic construct.
AUTHORS artificial sequence.
1 (bases 1 to 1071)
Bellenson,D., Smith,D., Lancet,D., Glusman,G., Fuchs,T. and
Yanal,I.
TITLE Olfactory receptor sequences
JOURNAL Patent: WO 0127158-A 242 19-APR-2001;
Discentis (US) ; YEDA RESEARCH AND DEVELOPMENT COMPANY, LTD. (IL)
FEATURES
source
1..1071
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="(H3891 nucleotide)"
BASE COUNT 277 a 248 c 201 g 345 t
ORIGIN

```

```

Query Match 77.0%; Score 726.2; DB 6; Length 1071;
Best Local Similarity 86.3%; Pred. No. 1.2e-195;
Matches 814; Conservative 0; Mismatches 128; Indels 1; Gaps 1;

QY 2 atgaatgggttaaatgacgacatcatacaggagttatctctgctgggttcctcaagtcga 61
DB 1 ATGAAATGGGTAAATAAGAGTGCACAGAGATTCTGTTAGATTCTCAGATCAA 60
QY 62 cctggagctgagttcaccctctgctctctctgattctcttaccactgagccacttt 121
DB 61 CCATGGCTAAGATTCACCCCTTTGTATGTCTTCTTCATATCTGACATCTTT 120
QY 122 ggaacatgcacattatctagtagtcagcctgagacacaaacttaacccccatgat 181
DB 121 GGCAATGTGACAAATATTTGTTGTGCATAGTGGATTTCAAACCTCCACACCCCTATGTAC 180
QY 182 tttttctcccaatctatactactctgagctctgtttaaaccacatgtlaagtcacca 241
DB 181 TTTTCTTGTGCAATCTCTACCTCGGACCTTGTGATCAACAAGTACAGTTCCACAA 240
QY 242 atgctaaatcatatgcagacatcagaagaagtaacagttatcgtgctgtgagccag 301
DB 241 ATGCTGTAAACATATGCAACACAGGAAGTATGATGTGCTGTGCTGTGCCCCAG 300
QY 302 ctttcaatctctgagccttgagggtactgaatctctctcctgscgtcatgtcctt 361
DB 301 CTTTTCATTTTCTGCGGCTTGGGTTCCACAGAAATGTTCTCCGCGCGCATGTGCTT 360
QY 362 gattggttttagctatttttgagcctctcattactacattacattacagaccagagctc 421
DB 361 GATAGGTTTATGCTATTTTGTGCGCTCTCATTTACTCAATATATACAGCAAGGCTC 420
QY 422 tggctcagcttgcagctgacatcctggtgttactggttttagtaactagagtggttct 481
DB 421 TGGCTTCAGTTGGCAGCTGATCCTGATTAAGTGGCTTTGACAAATTCAGATTACATCC 480
QY 482 accctgaactctcaagctgcacactctgtgaccctcatgtatgatatgaactctctgtga 541
DB 481 ACCTGACACTTAAAGATGCCACTGTGTGTACAAAGAAAGATGATCATCTTCTGTGAA 540
QY 542 gtccctgacactgtcgaagtatctgtgtgagagaacaaatgagggctggaacttc 601
DB 541 GTCCCTGCTGTGCTCAAGTTGTCTGTGTGACACACAAATGAGGCTGAACATAATTC 600
QY 602 ctgtcagtagactctctcatataacccccagacactacactctataataatgtctt 661
DB 601 TTTCATGAGTGTGATTTCTTAAATACCGGTAGCACTCATCTTAATATGATGCTTT 660

```

```

QY 662 atgtccgagcagfatatgagatatacagtcgtcgtgaaggtgcacaaaagcattgggaca 721
DB 661 ATGTGTCAACAGAGTGTGAGAAATCACTGTGTGAAGTCAACAGAAAGCATTTGGGACA 720
QY 722 tgtgttcccatctaatatgtgtgtctcttttatagtacagcgtctctgtgtactg 781
DB 721 TGTGTCTCCATCATTAATTTGTGTGTCACTTTTATGTGTACAGCTATCTCATGTACCTG 780
QY 782 caaccaccttcgcccagctcacaaggaccagaagaatgttctctctctatggaatc 841
DB 781 CAACCACCTTCACCCAGCTCCAAAGACGGGGAAGATGTTCTCTTCTGTGTGATC 840
QY 842 atggaccacatgctgtaatcccttatataatacactttagagacaagaaggtgaaggaagc 901
DB 841 ATTGACACCAFCGCGATTCCTCATATATACATTATGGAACAAAGAGTGAAGGAGCC 900
QY 902 tttaaaaggttggctgc-aagagctctcttaatacaagaataa 943
DB 901 TTTAAAGGTTGGTTCGCAAGAGTCTTATATCAAGAAATAA 943

```

```

RESULT 11
AX242075 1071 bp DNA Linear PAT 26-SEP-2001
LOCUS Sequence 823 from Patent WO0127158.
DEFINITION AX242075
ACCESSION AX242075
VERSION AX242075.1 GI:15798950
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
REFERENCE artificial sequence.
1 (bases 1 to 1071)
AUTHORS Bellenson,D., Smith,D., Lancet,D., Glusman,G., Fuchs,T. and
Yanal,I.
TITLE Olfactory receptor sequences
JOURNAL Patent: WO 0127158-A 823 19-APR-2001;
Discentis (US) ; YEDA RESEARCH AND DEVELOPMENT COMPANY, LTD. (IL)
FEATURES
source
1..1071
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="(H389673 nucleotide)"
BASE COUNT 277 a 248 c 201 g 345 t
ORIGIN

```

```

Query Match 77.0%; Score 726.2; DB 6; Length 1071;
Best Local Similarity 86.3%; Pred. No. 1.2e-195;
Matches 814; Conservative 0; Mismatches 128; Indels 1; Gaps 1;

QY 2 atgaatgggttaaatgacgacatcatacaggagttatctctgctgggttcctcaagtcga 61
DB 1 ATGAAATGGGTAAATAAGAGTGCACAGAGATTCTGTTAGATTCTCAGATCAA 60
QY 62 ccttgctgagttcaccctctgctctctctgattctcttaccactgagccacttt 121
DB 61 CCATGGCTAAGATTCACCCCTTTGTATGTCTTCTTCATATCTGACATCTTT 120
QY 122 ggaacatgcacattatctagtagtcagcctgagacacaaacttaacccccatgat 181
DB 121 GGCAATGTGACAAATATTTGTTGTGCATAGTGGATTTCAAACCTCCACACCCCTATGTAC 180
QY 182 tttttctcccaatctatactactcctgagctctgttataccacatgtatgaactccaca 241
DB 181 TTTTCTTGTGCAATCTCTACCTCGGACCTTGTGATATCCACAAAGTACAGTTCCACA 240
QY 242 atcctgaatattatgacagacatcagaagaagtaacagttatcgtgtgctgtgagccag 301
DB 241 ATGCTGTAAACATATGCAACACAGGAAGTATATGATGTGTGCTGTGCCCCAG 300
QY 302 ctttcaatctctgagccttgagggtactgaatctctctcctgscgtcatgtcctt 361

```

Db 301 CTTTTCATTTTCTGCGCTTGGGTTCCACAGATGTCCTCTCTGCGCGTCATGTCCTT 360
Qy 362 gattggtttgtatgataattgtcgccctccatctactcagatcatcatgacacagactc 421
Db 361 GATAGGTTTGTAGCTATTTGTGCGCTCTTCATTAATTAATGATGACACAGAGCTTC 420
Qy 422 tgcctcagttgagctcagctcagctcagctcagctcagctcagctcagctcagctcagctc 481
Db 421 TGCCTTCAGTTGCGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTC 480
Qy 482 accctgacctcagctcagctcagctcagctcagctcagctcagctcagctcagctcagctc 541
Db 481 ACCTGACACATTAAGATGCGACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
Qy 542 gtccctgacgtcagctcagctcagctcagctcagctcagctcagctcagctcagctcagctc 601
Db 541 GTCCCTGCTCTGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTC 600
Qy 602 ctgtgagtgagctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 661
Db 601 TTTATATAGTGTGCTATTCCTCTTAATACCCGTCACACTCATTCCTTAATGCTATTCCTTATTC 660
Qy 662 attgtccgacgagtgatgag 721
Db 661 ATTGTCACAGGAGGTGTGAAGATCCAGTCTGCTGGAAGGTCAACGAAGCATTTGGGACA 720
Qy 722 tctgttcccatcctaattgt 781
Db 721 TGTGGCTCCCATCTAATTTGT 780
Qy 782 caaccacgttcgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgac 841
Db 781 CAACGACCTTCACCCAGCTCCCAAGACCGGGGAAGATGTTCTCTCTCTGTGGAATC 840
Qy 842 atgtcacccatgctgataccctcctatatacactcagagacagagagagagagagagagagag 901
Db 841 ATGTCACCATGCTGATCCCTTATATATACACTTAGAGACAAAGAGGTAAAGAGAGCC 900
Qy 902 tttaaaaggtgtgtgc-aagagctcttctaatacaagaataa 943
Db 901 TTTTAAAGGTTGGTTCACAAAGCTCTTAAATCAAGAAATAA 943

RESULT 12
HSA302584 1074 bp DNA linear PRI 06-JAN-2001
LOCUS HSA302584
DEFINITION Homo sapiens 6M1-10*01 gene for olfactory receptor, cell line
BM28.7.
ACCESSION AJ302584
VERSION AJ302584.1 GI:12054392
KEYWORDS 6M1-10*01 gene; olfactory receptor.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1074)
AUTHORS Ziegler,A., Ehlers,A., Forbes,S.A., Trowsdale,J., Volz,A.,
Uchanska-Ziegler,B., Volz,A., Younger,R. and Beck,S.
TITLE Polymorphic olfactory receptor genes and HLA loci constitute
extended haplotypes
JOURNAL (in) Kasahara,M. (Ed.):
MAJOR HISTOCOMPATIBILITY COMPLEX-EVOLUTION, STRUCTURE, AND
FUNCTION: 110-130;
Springer-Verlag, Tokyo, Japan (2000)
REFERENCE 2 (bases 1 to 1074)
AUTHORS Ziegler,A., Ehlers,A., Forbes,S.A., Trowsdale,J., Volz,A.,
Younger,R.M. and Beck,S.
TITLE Polymorphisms in olfactory receptor genes: a cautionary note
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 1074)
AUTHORS Younger,R.M., Amdou,C., Bethel,G., Ehlers,A., Fischer,Lindahl,K.,
Forbes,S.A., Horton,R., Milne,S., Mungall,A.J., Trowsdale,J.,
Volz,A., Ziegler,A. and Beck,S.

TITLE Characterisation of clustered MHC-linked Olfactory Receptor Genes
JOURNAL In Human and Mouse
REFERENCE Unpublished
AUTHORS 4 (bases 1 to 1074)
Ehlers,A., Beck,S., Forbes,S.A., Trowsdale,J., Volz,A., Younger,R.
and Ziegler,A.
TITLE MHC-linked olfactory receptor loci exhibit polymorphism and
contribute to extended HLA/OR-haplotypes
JOURNAL Unpublished
REFERENCE 5 (bases 1 to 1074)
AUTHORS Ehlers,A.
TITLE Direct Submission
JOURNAL Submitted (06-NOV-2000) Ehlers A., Institut fuer Immunogenetik,
Universitaetsklinikum Charite, Spandauer Damm 130, 14050 Berlin,
GERMANY

FEATURES
source Location/Qualifiers
1..1074
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="6"
/cell_line="BM28.7"
1..1074
/gene="6M1-10*01"
1..1074
/gene="6M1-10*01"
CDS
1..1074
/gene="6M1-10*01"
codon_start=1
/product="olfactory receptor"
/protein_id="CAC20504.1"
/db_xref="GI:12054393"
translation="NMNWKSVDFEILVPSDPMLEPPVNFLEFSLTIFGNLT
ITLVNHPDKLTIPKVFPLSNLIDICTVSTVOMLVNCTKPKVSYGGCVOLF
ITLFGSTFECILLAVMCDRFAITRPLHYSITIMORLCFQILAAWSISGFSNVLQ
TWTLMPLCGHKEVDHFCFCEVPALIKLSVDPTTAEAELFISVLLIPVLLISTY
AFIVQVILRIQSAEGQRKAFGCGSHLIVSLFETALISMVLOPPSPSKDQKMWSL
FCGIIAPVMLNPLITLIRNKEVKAERKRLVAKSLINDEIRNMOMISFAKDTVLTILNF
SASCPLEPVTINCNLPORKP"

BASE COUNT 278 a 248 c 202 g 346 t
ORIGIN

Query Match 77.0%; Score 726.2; DB 9; Length 1074;
Best Local Similarity 86.3%; Pred. No. 1.2e-195;
Matches 814; Conservative 0; Mismatches 128; Indels 1; Gaps 1;

Qy 2 atgaattgggttaaattgacagcatalcacaagagattattctgtcgtgggttcagaatcga 61
Db 1 ATGAATTGGGTAAATTAAGAGTGCACAGAGAGTCAATCTGTTAGTTTCTCAGATCAA 60
Qy 62 ccttgctgagagttccactcctgtgtgtctcttgattcttaacatgagaccatctt 121
Db 61 CCATGGCTAGAGATCCACCCTTGGATGTTCTGTTTCTATATCTTGACAAATCTTT 120
Qy 122 ggcacatgcacattattctagtgtcagcgctgacacacaaactcatcacccatgfat 181
Db 121 GGCATCTGCATTAATTTCTTGTGTCAAGTGTGATTTCAAACTCCACACCCCTATGTAC 180
Qy 182 tttttcttaccatcatcactcctgagactctgtttacacacacatgtaacagtcaccca 241
Db 181 TTTTCTTGTAGCAATCTCTCAGCTGAGACCTTGTGTAATACACAGTACAGTTCCACAA 240
Qy 242 atcgtcgttaattatgacagcagcaagaagtaatacagttatcgtgtgctgtgagcccg 301
Db 241 ATCGTGTGAACATATGCAACACACAGAAAGTAATCAGTTATGTGCTGTGGCCACG 300
Qy 302 cttcataattctggccttgagggtgactggaatactcttctcgtggcggtcatgctctt 361
Db 301 CTTTTCATTTTCTGCGCTTGGGTTCCACAGATGTCCTCTCTGCGCGTCATGTCCTT 360
Qy 362 gattggtttgtatgataattgtcgccctccatctactcagatcatcatgacacagactc 421
Db 361 GATAGGTTTGTAGCTATTTGTGCGCTCTTCATTAATTAATGATGACACAGAGCTTC 420
Qy 422 tgcctcagttgagctcagctcagctcagctcagctcagctcagctcagctcagctcagctc 481

DB	421	TCCTTCGATTGGACGCTCATCTCGAGATTAGGCTTTAGCAATTCACATATTACAGTCC	480
OY	482	accctgaactctccagctgcgaactctgtgacccctatgtatagatacttctctgtga	541
DB	481	ACCTGGACACTTAAAGATGGCACTGTGTGTCACAAAGATGGATCATTCTTGTGTGA	540
OY	542	gtccctgcactgtcaagtaactctgtgttgtagacacagacaagttagctgaacttc	601
DB	541	GTCCCTGCTCTCTCAAGTGTGTCTGTGTGACACACAGCAAGCAATGAGCTGACATTC	600
OY	602	cttgtaagtagctcttcatactbaabaccctgcgaactcaccctataatcatatgcttt	661
DB	601	TTCAATCAGTGTCTATTCTCTTCTATATACCCGGACACTATCTTATATGATGCTTTT	660
OY	662	attgtcccgagagatattggagatacagctctgttgaaggtcgaacaaacatttggaca	721
DB	661	ATTGTCCAAGCAGTGTGGAAATCCAGTGTGCTGAAGGTCAACGAAAGCAATTTGGACA	720
OY	722	tgtagttcccatctaattgtgtgtgtctcttttataagacagccgtctctgtacccg	781
DB	721	TGTGCTCCCATCTATATTGTGTGTGTCATTTTATATAGTACAGCTATCTCATGTACGTG	780
OY	782	caaccactctggcccaagctcccaagacccaagaagaatgtttctctctctatgaaac	841
DB	781	CACACACCTTCAACCAAGCTCCAAAGCCGGGAAAGATGTTCTCTCTGTGTGAATC	840
OY	842	attgacaccatctgaatcccttatatactatagacaggaagagtaagaagaagc	901
DB	841	ATTGCACCCATCTGTGATCCCTTATATATACACTTAGACAAAGAGGTTAAAGGAAAGCC	900
OY	902	tttaaaagtggtgtgtgc-aagaagctcttctatcaagaata	943
DB	901	TTTAAAGGTGTGTGCAAAAGATCTTCTTATCAAGAAATPA	943

RESULT 13

HS302585

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

1074 bp

DNA

linear

PRI 06-JAN-2001

Homo sapiens

6M1-10*01 gene for olfactory receptor, cell line

BMJ 9, .

AJ302585

AJ302585.1

GI:12054394

6M1-10*01 gene; olfactory receptor.

human

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrate; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 1074)

Ziegler,A., Ehlers,A., Forbes,S.A., Trowsdale,J., and Beck,S.

uchanaka-Ziegler,B., Volz,A., Younger,R. and Beck,S.

Polymorphic olfactory receptor genes and HLA loci constitute extended haplotypes

(1) Kasahara,M. (Ed.);

MAJOR HISTOCOMPATIBILITY COMPLEX-EVOLUTION, STRUCTURE, AND FUNCTION: 110-130;

Springer-Verlag, Tokyo, Japan (2000)

2 (bases 1 to 1074)

Ziegler,A., Ehlers,A., Forbes,S.A., Trowsdale,J., Volz,A., Younger,R.M. and Beck,S.

Polymerphisms in olfactory receptor genes: a cautionary note

unpublished

3 (bases 1 to 1074)

Younger,R.M., Amadou,C., Bethel,G., Ehlers,A., Fischer Lindahl,K., Forbes,S.A., Horton,R., Milne,S., Mungall,A.J., Trowsdale,J., Volz,A., Ziegler,A. and Beck,S.

Characterisation of clustered MHC-linked Olfactory Receptor Genes in Human and Mouse

Unpublished

4 (bases 1 to 1074)

Ehlers,A., Beck,S., Forbes,S.A., Trowsdale,J., Volz,A., Younger,R. and Ziegler,A.

MHC-linked olfactory receptor loci exhibit polymorphism and

JOURNAL REFERENCE AUTHORS TITLE JOURNAL	contribute to extended HLA/OR-haplotypes 5 (bases 1 to 1074) Ehlers, A. Direct Submission Submitted (06-NOV-2000) Ehlers A., Institut fuer Immunogenetik, Universitaetsklinikum Charite, Spandauer Damm 130, 14050 Berlin, GERMANY
FEATURES	Location/Qualifiers
Source	1..1074 /organism="Homo sapiens" /db_xref="taxon:9606" /chromosome="6" /cell_line="BM19.7" 1..1074 /gene="6M1-10*01" 1..1074 /gene="6M1-10*01" /codon_start=1 /product="olfactory receptor" /protein_id="CAC20505.1" /db_xref="GI:12054395" /translation="MMRVNKSVPQEFLLVPSDDPWLIEPFVLEFSYLITFGNLT IIIVSHDEKHTPMYFFPLSLSLDLICYTTSYPMQLVNINCRKISYGGCAOLF IFLALSGTECLLALVWCFDEFRVAICRPDHSIIMHQLCFLAASVISGNSVLOS TWLTKMPCGKHEVDHFECEFPALALICSDVTANAELEFFISVLFILIPVLLISY AFIVQAVLRIOASGORKARFCGTCGSHLIVSLFPGTALISMVLPSPSSKDRGKWSL FCGJIALPMLNDLIATLRNKEVKEAFKFLKVALSLNDEIRNMQMSIFAKDVLVLTVMF SASCFIVETIENCNLPORKEP"
gene	
CDS	
BASE COUNT	278 a 248 c 202 g 346 t
ORIGIN	
Query Match	77.0%; Score 726.2; DB 9; Length 1074;
Best Local Similarity	86.3%; Pred. No. 1.2e-195;
Matches 814; Conservative	0; Mismatches 128; Indels 1; Gaps 1;
Oy	2 atgaattggtgtaaatgagaagcatcatataagaagttattctgctgggtcttcagatcga 61
Db	1 ATGAATTTGGTAATAAAGATGTCCCAAGAGAGGTCATTCGTGTAAGTTTCCACAGATCAA 60
Oy	62 ccttgctgaggttcceagctccttctggtctctctgtatcttcttaccagctgacacatc 121
Db	61 CCATGGCTAGAGATTCCACCCCTTTGTGATGTTTCTGTTTCTCTATATCTTGACAACTTT 120
Oy	122 ggcaactgaccatattcttagtgtgcagcctgagacccaacttcataccccaatgtat 181
Db	121 GGCAATCTCGAACAATATATCTGTGTGACATGTGGATGTTTCAAACTCCACACCCCTATGTAC 180
Oy	182 tttttcttaccacatcatcaactcctctgtagctcttcttaccacaaatgtacagctccaaa 241
Db	181 TTTTTCCTTAGCAATCTCTCACTCCTGGAACCTTTTGCTATACACAAAGTACAGTTCACAA 240
Oy	242 atgctagtaaaattatgcaagcatcagaagaatlaactcagttatcgttggctgttagccag 301
Db	241 ATGCTGTAATAACATATGTGACAAACCCAGAAATTAACATATATGTGAGGCTGTGGGCCAG 300
Oy	302 ctttccatattcttggccttggggctgtactgaatactcttcctcgccgtcatgtccctt 361
Db	301 CTTTTCATTTTCTCGGCTGTGGCTGTCCACAGAAATGTCTTCTCGGCCGTATGTGCTTT 360
Oy	362 gatgttgatgactattctgctcgagccttcacatctactcaagtattcatcaccagagagac 421
Db	361 GATAGTGTTAGAGCATTTGTGTGGCTCTCCATTAATCACTCAATTAATCAATGACACAGAGTC 420
Oy	422 tgcctccagttgctgcaagctgcatacctctggtttactcgtgttttagtlaactcagtggtc 481
Db	421 TGCTTCCAGTTGGGAGCTGCATCCGATTTAGTGGCTTTTACCAATTCAGTATACAGTCC 480
Oy	482 accctgactctccagctgcagctctgtgaaaccccatgtgataagataactctctctgtgaa 541
Db	481 ACCTGGAACCTTAAGATGCCACTGTGTGTACAAAGAAAGATGATCACTCTTCTGTGAA 540

QY	542	gtccctgaacgctcaaaattatcttggtttgaagacaacagaacaaatgagcgtgaactattc	601
Db	541	gtccctgctctgctcmaattggttcctcgittgacacaaacagaaatgagcgtgaactattc	600
QY	602	cttgacatgagcgtcttcacatcaatatacccccgcagacatcatcattcatatgctttt	661
Db	601	tttcacactgagctgctatgcttcttatactaccggacacacatcattatgcttctttt	660
QY	662	attgtcgaagcaagatctgagatatacagatcgtctgtaaggttcgacaaaagaatttggaca	721
Db	661	attgtccaaagcagtgcttgagaatccagctcgtctgtaaggttcacaaagacatttgggaca	720
QY	722	tgtgtgtccatctaatgtgtgtgtgtgtgtgtgtgtttttatagtaagccgtctctgtgacctg	781
Db	721	tgtgtgtccatctaatgtgtgtgtgtgtgtgtgtgtttttatagtaagccgtctctgtgacctg	780
QY	782	caaccacacttcgcccagcgtcccaaggaacgaagaagatggtttctctctctcatgtaalc	841
Db	781	caaccacacttcacaccagctccaaagaccgggaaagatggtttctctctctctgctgaatc	840
QY	842	attgcacccatgctgtaatcccttatatacatcattaggaacaaagaggtaaaggaagc	901
Db	841	attgcacccatgctgtaatcccttatatacatcattaggaacaaagaggtaaaggaagc	900
QY	902	tttaaaaggtgtgtgtgc-aagagttcttcttaatacagaataa	943
Db	901	ttttaaaggtgtgtgtgc-aagagttcttcttaatacagaataa	943
RESULT	14		
LOCUS	HSA302586	1074 bp	DNA linear PRI 06-JAN-2001
DEFINITION	Homo sapiens 6M1-10*01 gene for olfactory receptor, cell line LG2.		
ACCESSION	AJ302586		
VERSION	AJ302586.1	GI:12054396	
KEYWORDS	6M1-10*01 gene; olfactory receptor.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
AUTHORS	Ziegler, A., Ehlers, A., Forbes, S.A., Trowsdale, J., and Beck, S.		
TITLE	Polymorphic olfactory receptor genes and HLA loci constitute extended haplotypes		
JOURNAL	(In) Kasahara, M. (Ed.): MAJOR HISTOCOMPATIBILITY COMPLEX-EVOLUTION, STRUCTURE, AND FUNCTION: 110-130; Springer-Verlag, Tokyo, Japan (2000)		
REFERENCE	2 (bases 1 to 1074)		
AUTHORS	Ziegler, A., Ehlers, A., Forbes, S.A., Trowsdale, J., Voltz, A., Younger, R.M. and Beck, S.		
TITLE	Polymorphisms in olfactory receptor genes: a cautionary note		
JOURNAL	Unpublished		
REFERENCE	3 (bases 1 to 1074)		
AUTHORS	Younger, R.M., Anadou, C., Bethel, G., Ehlers, A., Fischer Lindahl, K., Forbes, S.A., Horton, R., Milne, S., Mungall, A.J., Trowsdale, J., Voltz, A., Ziegler, A. and Beck, S.		
TITLE	Characterisation of clustered MHC-linked Olfactory Receptor Genes in Human and Mouse		
JOURNAL	Unpublished		
REFERENCE	4 (bases 1 to 1074)		
AUTHORS	Ehlers, A., Beck, S., Forbes, S.A., Trowsdale, J., Voltz, A., Younger, R. and Ziegler, A.		
TITLE	MHC-linked olfactory receptor loci exhibit polymorphism and contribute to extended HLA/OR-haplotypes		
JOURNAL	Unpublished		
REFERENCE	5 (bases 1 to 1074)		
AUTHORS	Ehlers, A.		
TITLE	Direct Submission		
JOURNAL	Submitted (06-NOV-2000) Ehlers A., Institut fuer Immunogenetik, Universitaetsklinikum Charite, Spandauer Damm 130, 14050 Berlin, GERMANY		

FEATURES		Location/Qualifiers
source		1..1074 /organism="Homo sapiens" /db_xref="taxon:9606" /chromosome="6" /cell_line="J62" 1..1074 /gene="6M1-10+01" 1..1074 /gene="6M1-10+01" /gene="6M1-10+01" /codon_start=1 /product="olfactory receptor" /protein_id="CAC30506.1" /db_xref="GI:12054397" /translation="MNNVNSVDPDEFILVFSDDPMLIEIPVEWELSYILITFGNLT ILIVSHVDFKLHTPMYFFSLNLSLDDCYTSTVPQMLVINCNRKVISGGVADQLE IFLALGSETECLILAVMCFEDRFLVACRPLHYSIIMHQRLCPLAASWISSEFNSVLOS TWMLKPLRGKHKEVDHFECECPALTKIKISVDPTNKELEFISVLPILIVTILISY AFIVQAVLRIDSABGQRKARGTGSCHLIVSYLPFGRAISMTIDPPSSSDRGKMSVL FCGLIAPMLNPILITLNRKEKFEAKFKLVAKSLNIGLRNMQIMSFKAIDVLYLTINE SASPIPIVITLIENCNLPORKEP"
CDS		
BASE COUNT	278 a 248 c 202 g 346 t	
ORIGIN		
Query Match	77.0%; Score 726.2; DB 9; Length 1074;	
Best Local Similarity	86.3%; Pred. No. 1.2e-195;	
Matches	814; Conservative 0; Mismatches 128; Indels 1; Gaps 1;	
OY	2 atgaattggttaaatagacagcatcatatacagaagttattctgcgtggtttccagatga	61
DB	1 ATGAATTTGGGTAATAAAGTGTCCCAAGAGCATTCATTGTGTATGTCAGAACTCA	60
OY	62 ccttgagctggagttccagactccctgtgtgtctctctgattctctacacgtgacacttt	121
DB	61 CCATGGCTAGAGATTTCCACCCCTTTGTGATGTTTCTGTTTCCATAATCTTGACAAATCTTT	120
OY	122 ggcacatctgacattattctatagttgcacagccttgagacacaaactcatataccccaatgtat	181
DB	121 GGCATCTGACAAATAATTTCTGTGTGCACATGATGGATTTCAAACTCCACACCCCTATGTATC	180
OY	182 ttttttcttacaacatcatcaactcctctgtaactctgtttacacacaatgtacagttccacaa	241
DB	181 TTTTTCCTTTAGCAAACTCTCACTCCTCGAGACCTTTTCTCTTACCAACAGTACAGTTTCCACAA	240
OY	242 atgcagtaaatatttatgcagcatcagagaaagaatcagttatcgtggctgtgtagccag	301
DB	241 ATGCTGGTAACATATGACACACACGGAAGTAATACAGTTATGGTGGCTGTGTGGCCACG	300
OY	302 ctcttcacatatctctgaccttgaggctactagtaatactcttccctctgagccgtcaatgctctt	361
DB	301 CTTTTCATTTTTCCTGGCCCTGTGGCTTCCAAAGATGTCTTCTCCTGGCCGTATGTGCTTT	360
OY	362 gattggtttgttagctatttcttcgcgcctctccatcaactcaagtatatacagcagaagactc	421
DB	361 GATTAGTTTGTATGTGATTTGTGGCCCTCTCCATTCCTCAATATATATCACCACGAGGGCT	420
OY	422 tgcctccagttggagagctgcgaatcctctgggttttaactggttttagtaactagtggtgtcct	481
DB	421 TGTCTTCCAGTTTGGCAGCTGCATCTGGATTTATGTGCTTTTAGCAATTCAGTATTACAGTCC	480
OY	482 accctgcactctccagactgcgaactctgttgaacccctatgtgatagatcaacttctctgtgaa	541
DB	481 ACGTGGACACTTAAATATGCACTGATGTGTGTGCACAAAGATGATGATCATTCTTCTGTGA	540
OY	542 gtccctgcagctcgaattcatctctgtgttttgagacacagcaaatgagctgtaactatcc	601
DB	541 GTCCCTGCTGTGCTCAAGATTGTGCTGTGTGTGCACAAAGAGAGAGCTGAATCTATTC	600
OY	602 ctgtgcagttgagcttctcatatacaccctggaacacatccatatacgtatgctttt	661
DB	601 TTTCATCATGTGCTATTCCTTCTAATATACCGGACATCATCTTATATATGATATCTTTT	660

OY	662	atgfcgcgaagcagttattagaataacagctcgtagaagtgcgacaanaagaacttgggaca	721
Db	661	ATTGTCCAAAGCACTGTTTGGAATTCACGTCGTGTGAAGGCACAAGAAGCATTTGGACA	720
OY	722	tgtgttcaccatcatattgtgtgtcctctttttaatagtaacagccgctctgtgacctg	781
Db	721	TGTGCTCCGANTCTAATTTGTGTGTACTTTTTTATGTGTAGACGTAATCTCATTGACTCG	780
OY	782	caaacacacctgcgccagctccaaggaccaagaaagatgtttctctctctctaagaaac	841
Db	781	CACACCACCTTCACCCAGCTCCAAAAGACCGGGGAAAAAGATGGTTTCCTCTCTGTGGAATC	840
OY	842	atttgaccacatgccttaatcccccttatataoacttagaacaagaaggtaaagaaagc	901
Db	841	ATTGCAACCCATGCTAAATCCCCTTATATATACATACTTAGAACATAAAGAGTAAGAAGGCC	900
OY	902	tttaaaaggttgtgtgc-aagaagctctcttaatacaagaataa	943
Db	901	TTTAAAGGTTCGTTGCCAAAGAGCTTCTTATATCAAGAATAA	943
RESULT 15			
HSA302587		1074 bp	DNA linear PRI 06-JAN-2001
LOCUS			
DEFINITION	Homo sapiens 6M1-10*01 gene for olfactory receptor, cell 11ne		
ACCESSION	H2LCL		
VERSION	AJ302587		
KEYWORDS	AJ302587.1 GI:12054398		
SOURCE	6M1-10*01 gene; olfactory receptor.		
ORGANISM	human.		
	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
REFERENCE	1 (bases 1 to 1074)		
AUTHORS	Ziegler,A., Ehlers,A., Forbes,S.A., Trowsdale,J., Uchanska-Ziegler,B., Volz,A., Younger,R. and Beck,S. Polymorphic olfactory receptor genes and HLA loci constitute extended haplotypes (in) Kasahara,M. (Ed.). MAJOR HISTOCOMPATIBILITY COMPLEX-EVOLUTION, STRUCTURE, AND FUNCTION: 110-130; Springer-Verlag, Tokyo, Japan (2000)		
JOURNAL	2 (bases 1 to 1074)		
REFERENCE	Younger,R.M. and Beck,S.		
AUTHORS	Polymorphisms in olfactory receptor genes: a cautionary note Unpublished 3 (bases 1 to 1074)		
JOURNAL	Younger,R.M., Amadou,C., Bethel,G., Ehlers,A., Fischer Lindahl,K., Forbes,S.A., Horton,R., Maline,S., Mungall,A.J., Trowsdale,J., Volz,A., Ziegler,A. and Beck,S. Characterisation of clustered MHC-linked Olfactory Receptor Genes in Human and Mouse Unpublished 4 (bases 1 to 1074)		
REFERENCE	Ehlers,A., Beck,S., Forbes,S.A., Trowsdale,J., Volz,A., Younger,R.		
AUTHORS	and Ziegler,A.		
JOURNAL	MHC-linked olfactory receptor loci exhibit polymorphism and contribute to extended HLA/OR-haplotypes Unpublished 5 (bases 1 to 1074)		
REFERENCE	Ehlers,A.		
AUTHORS	Direct Submission Submitted (06-NOV-2000) Ehlers A., Institut fuer Immunogenetik, Universitaetsklinikum Charite, Spandauer Damm 130, 14050 Berlin, GERMANY		
FEATURES			
source	Location/Qualifiers		
	1..1074		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/chromosome="6"		
	/cell_line="H2LCL"		
	1..1074		
gene			

CDS	/gene="6M1-10+01 1..1074 /gene="6M1-10+01" /codon_start=1 /product="Olfactory receptor" /protein_id="CAC20507.1" /db_xref="GI:12054399" /translation="MNNVNSVPEPFLLPFSODPLPEFPWFLESLITFGNLT ILIVSHDEFLKHPMPVFFLSLLDLCTVTSTPQWLNNCRKIVSGVADLP IFLAVSGTEFLIAVMCFDRFVALICRPLHSIIMHOCQLQLANRKISGNSVLQS TWTKMPLGCHKEVDHFPECEPALKLSVDYTNMELEFIVLPIPVTLILISY AFIVQAVLRIOASGORKARCTGCSHLIVSLFPGTAIMSYLQPPSSSKDRKMVEL FCGIIAAMNLIYTLIRNKEFKRFLVAKSLNLNGEIRNMNMISFAKDIYLVTLTNE SASCPPIVETIENCNIPORKE"
BASE COUNT	278 a 248 c 202 g 346 t
ORIGIN	
Query Match	77.0%; Score 726.2; DB 9; Length 1074; Best Local Similarity 86.3%; Pred. No. 1.2e-195;
Matches 814: Conservative	0; Mismatches 128; Indels 1; Gaps 1;
Oy	2 atgaattggtaaaagacacatcacatagaagattatctgctgggtcttcagatcga 61
Db	1 ATGATTTGGGTAATTAAGAGTGTCCACAGAGGTTCATCTGTATGTTTCTGCAGATCAA 60
Oy	62 ccttgagctggagattccacacctggtgctctcttgattcttaccactgacacctt 121
Db	61 CCATGCTGAGAGATTCCACCCTTTGTGATGTTTCTGTTTCTATATCTTGACAACTT 120
Oy	122 ggcaatctgacattatcttagtgcacagccctggacacaaactcataccoccatg 181
Db	121 GGCAATCTGACAAATAATCTTGTGTGCACATGTGGATTTTCAAACTCCACACCCTATGTGAC 180
Oy	182 tttttcttaccacatcatcacacctctggtatcttttliacacacatgacatccacaa 241
Db	181 TTTTTCCTTAGCAATCTCTCACTCCTCGACCTTTCGTATACCAACAAGTACAGTTCCACAA 240
Oy	242 atgctagtaaaatttatgacagcatcagaagaagaaatcagttacgtgctgtgtagccag 301
Db	241 ATGCTGTGTAACATATGTCACACACACGGAAGAATAATCATGTATGTGGCTGTGTGGCCAG 300
Oy	302 ctttccatattcttgacctgggggtactgataatcttccctggccgctcagtcctt 361
Db	301 CTTTTCATTTTCCCTGGCCCTTGGGTTCCACAGATGTCCTTCCTGGCCGCTATGTGCTTT 360
Oy	362 gatggtcttgtagctattctgtgcgctctccatccatcagttatcatgacacagaagact 421
Db	361 GATGAGTTTGTAGCTATTTGTGTGCGCTCTCCATTACTAATTAATCATCATCACCAGAAGCTC 420
Oy	422 tgcctccagattggcagctgcacatcctgggttactggttctttagtlaactcagtgtygtct 481
Db	421 TGCTTCAGATTGGCAGCTGCATCCTGGATTTAGTGGCTTTAGCAATTCAGTATTTACAGTCC 480
Oy	482 accctgacctccagctgcacacctctgtgaacctatgtgatagatcaacttctctgtgaa 541
Db	481 AACTGGACACTTAATCATCCACTGTGTGTGTCACAAAGAAGTGATCACCTTCTTGCTCAA 540
Oy	542 gtccctgacagctcacaagttaatctgtgtgtgagacaaagaagaatgagcgaactatcc 601
Db	541 GTCCCTGCTGTGCTCAAACTGTGTGTGTGTGCACAAACGAAGCAATGAGGCTGAACATAATTC 600
Oy	602 ctgtgcagtgagctcttcacatctaaataccctgacacatcctatatactatgctt 661
Db	601 TTTCATCATGTGTGCTATTCCTCTCTATATACCGGAGACATCATCTTATATGATAGTCTTTT 660
Oy	662 atgtgcagagagtatgtgagatacagtcgtcgaaggtcgacaanaaagcatltggagaca 721
Db	661 ATTGTCCAAGACAGTGTGAGAAATCCAGTCTGCTGAAGGTCAACGAAGAAGCATTTGGGACA 720
Oy	722 tgtgtgtcccatcgaatgtgtgtgtctctttttatagtaagccgctctcgtgtactgt 781
Db	721 TGTGTGCTCCCATATTAATGTGTGTGTGCTATTTTATATGAGTATAGCTATTCATGTACTGTG 780

QY 782 caaccacccctcgcccaagctccaaggaagaagatggttctctctctatggaatc 841
Db 781 CAACCACTTCACCCAGCTCCAAAGACCGGGGAAGAATGGTTCTCTCTCTGTGAATC 840
QY 842 atgcaccatgctgaaatcccttatatatataacttaggaacaagaggtaaaggaagc 901
Db 841 ATTGCACCCATGCTGAATCCCTTATATATACCTTAGGAACAAGAGTAAAGGAAGCC 900
QY 902 tttaaaagtggttgc-aagagctctctaatacaagaataa 943
Db 901 TTTAAAGGTTGTTGCAAGAGTCTTATATCAAGAAATAA 943

Search completed: August 25, 2002, 18:57:02
Job time: 14460 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd

OM nucleic - nucleic search, using sw model

Run on: August 25, 2002, 17:52:12 ; Search time 265.69 seconds
(without alignments)
6093.757 Million cell updates/sec

Title:	US-09-800-321A-3
Perfect score:	943
Sequence:	1 tatgatcggtaatgaca.....tcttctaaccagaataa 943

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

```
Searched:      1736436 seqs, 858457221 residues
Total number of hits satisfying chosen parameters: 3472872
```

```
Minimum DB seq length: 0
Maximum DB seq length: 20000000000
```

```
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
```

N_Geneseq_032802.*

1	/SIDS5/gcgdata/genseq/genseqn-emb1/NA1980.DAT.*
2	/SIDS5/gcgdata/genseq/genseqn-emb1/NA1981.DAT.*
3	/SIDS5/gcgdata/genseq/genseqn-emb1/NA1982.DAT.*
4	/SIDS5/gcgdata/genseq/genseqn-emb1/NA1983.DAT.*
5	/SIDS5/gcgdata/genseq/genseqn-emb1/NA1984.DAT.*
6	/SIDS5/gcgdata/genseq/genseqn-emb1/NA1986.DAT.*
7	/SIDS5/gcgdata/genseq/genseqn-emb1/NA1987.DAT.*
8	/SIDS5/gcgdata/genseq/genseqn-emb1/NA1988.DAT.*
9	/SIDS5/gcgdata/genseq/genseqn-emb1/NA1989.DAT.*
10	/SIDS5/gcgdata/genseq/genseqn-emb1/NA1990.DAT.*
11	/SIDS5/gcgdata/genseq/genseqn-emb1/NA1991.DAT.*
12	/SIDS5/gcgdata/genseq/genseqn-emb1/NA1992.DAT.*
13	/SIDS5/gcgdata/genseq/genseqn-emb1/NA1993.DAT.*
14	/SIDS5/gcgdata/genseq/genseqn-emb1/NA1994.DAT.*
15	/SIDS5/gcgdata/genseq/genseqn-emb1/NA1995.DAT.*
16	/SIDS5/gcgdata/genseq/genseqn-emb1/NA1996.DAT.*
17	/SIDS5/gcgdata/genseq/genseqn-emb1/NA1997.DAT.*
18	/SIDS5/gcgdata/genseq/genseqn-emb1/NA1998.DAT.*
19	/SIDS5/gcgdata/genseq/genseqn-emb1/NA1999.DAT.*
20	/SIDS5/gcgdata/genseq/genseqn-emb1/NA2000.DAT.*
21	/SIDS5/gcgdata/genseq/genseqn-emb1/NA2001.DAT.*
22	/SIDS5/gcgdata/genseq/genseqn-emb1/NA2001A.DAT.*
23	/SIDS5/gcgdata/genseq/genseqn-emb1/NA2001B.DAT.*
24	/SIDS5/gcgdata/genseq/genseqn-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No	Score	Query Match	Length	DB	ID	Description
1	943	100.0	943	22	AAD17016	G-protein coupled
2	940.4	99.7	993	22	AAD17015	G-protein coupled
3	938.8	99.6	942	22	AA508541	DNA encoding novel
4	938.8	99.6	1488	22	AA508542	ORF encoding novel
5	937.4	99.4	939	22	AAH32018	Human olfactory re
6	937.2	99.4	956	22	AAD17017	G-protein coupled
7	726.2	77.0	1071	22	AAH31659	Human olfactory re
8	726.2	77.0	1071	22	AAH32250	Human olfactory re
9	641.6	68.0	648	22	AAH32331	Human olfactory re

10	533	56.6	1442	21	AACT7475	Human OREF ORF3030
11	531	55.3	939	22	AAH33304	Human olfactory receptor neuron
12	403.6	42.8	939	22	AAAS42221	Human CDNA encoding
13	402	42.6	936	22	AAH31645	Human olfactory receptor
14	402	42.6	939	24	ABK16619	Human G-coupled receptor
15	371.6	39.4	1002	22	AAH32063	Human olfactory receptor
16	370	39.2	1005	22	AAAS42391	Human CDNA encoding
17	370	39.2	1065	22	AAH78086	Breast amplified G-protein coupled
18	364.2	38.6	1011	22	AAH32306	Human olfactory receptor
19	356.6	37.8	933	22	AAH33308	Human olfactory receptor
20	356	37.8	936	22	AAH32314	Human olfactory receptor
21	355.6	37.7	927	22	AAH33067	Human olfactory receptor
22	355.6	37.7	1053	22	AAAS42407	Human CDNA encoding
23	354	37.5	939	22	AAAS42428	Human CDNA encoding
24	352.4	37.4	936	22	AAH32045	Human olfactory receptor
25	350.2	37.1	1014	22	AAH32316	Human olfactory receptor
26	350.2	37.1	1440	21	AACT6560	Human OREF ORF2515
27	345	36.6	1035	22	AAH78083	Breast amplified G-protein coupled
28	343.2	36.4	936	22	AAH31632	Human olfactory receptor
29	343.2	36.4	1954	22	ABA57028	Human foetal liver
30	343.2	36.4	1954	22	ABA26532	Probe 45098 for gene
31	341.2	36.2	979	22	AAH32025	Human olfactory receptor
32	337.6	35.8	1013	23	AAAD17919	G-protein coupled receptor
33	336.4	35.7	1179	23	AAST77100	DNA encoding novel
34	336.4	35.7	1302	23	AAST77093	DNA encoding novel
35	335.8	35.6	909	22	AAH32043	Human olfactory receptor
36	334.2	35.4	1629	22	AAAD0980	Human novel 32164
37	329.8	35.0	947	22	AAAD1908	G-protein coupled
38	328.8	34.9	948	22	AAH31841	Human olfactory receptor
39	328.8	34.9	948	22	AAH32375	Human olfactory receptor
40	328.2	34.8	1002	22	AAH32238	Human olfactory receptor
41	326.4	34.6	932	22	AAAD17912	G-protein coupled
42	324	34.4	948	22	AAH32188	Human olfactory receptor
43	319.6	33.9	1351	22	AAH78085	Breast amplified G-protein coupled
44	318	33.7	957	22	AAH32072	Human olfactory receptor
45	318	33.7	957	22	AAH32088	Human olfactory receptor

ALIGNMENTS

RESULT	1
AAD17016	
ID	AAD17016 standard; DNA; 943 BP.

AC	AAD17016;
XX	
DT	29-NOV-2001 (first entry)

DE G-protein coupled-receptor (GPCR) NOV1b DNA.

	Key	Location/Qualifiers
FW	5'UTR	1
FT		/*tag= a
KW		G-protein coupled receptor; GPCR; NOVX-associated disorder; obesity;
KW		anorexia; wasting disorder; infection; cachexia; Parkinson's disease;
KW		Alzheimer's disease; immune disorder; neurodegenerative disease; cancer;
KW		anorectic; haematopeletic disorder; major histocompatibility complex;
KW		MHCII; noninsulin-dependent diabetes mellitus; NIDDM; bulimia; asthma;
KW		acute heart failure; hypotension; multiple sclerosis; hypertension;
KW		osteoporosis; Crohn's disease; mental retardation; dementia; allergy;
KW		angina pectoris; myocardial infarction; benign prostatic hypertrophy;
KW		psychotic disorder; neurological disorder; anxiety; schizophrenia;
KW		manic depression; delirium; Albright Hereditary Osteodystrophy; rickets;
KW		dentalorodipallidolusian atrophy; DRP1A; haemostatic; anticoagulant;
KW		autosomal dominant (2) acrocallosal syndrome; dyskinesia; neuroleptic;
KW		Huntington's disease; gilles de la Tourette syndrome; neuroprotective;
KW		nutropic; antibacterial; protozoacide; fungicide; cytostatic; NOV1b;
KW		vulnerary; cardiant; osteopathic; antianginal; anticidre; viuncide;
KW		human immunodeficiency virus; HIV; antidepressant; tranquilliser; ds.
XX		Unidentified.
OS		
XX		
FW		
FT		

FT	CDS	2..943	/*tag= b	
FT		2..115	/product= "GPCR NOV1b protein"	
FT	sig_peptide	2..115	/*tag= c	
FT	mat_peptide	116..940	/*tag= d	
FT		/product=	"Mature GPCR NOV1b protein"	
XX		WO20016746-A2.		
XX		13-SEP-2001.		
XX		05-MAR-2001; 2001WO-US07116.		
XX		03-MAR-2000; 2000US-0186606.		
XX		06-MAR-2000; 2000US-0187247.		
XX		06-MAR-2000; 2000US-0187248.		
XX		06-MAR-2000; 2000US-0187249.		
XX		06-MAR-2000; 2000US-0187250.		
XX		06-MAR-2000; 2000US-0187253.		
XX		06-MAR-2000; 2000US-0187295.		
XX		07-MAR-2000; 2000US-0187296.		
XX		21-JUL-2000; 2000US-0219854.		
XX		24-JUL-2000; 2000US-0220263.		
XX		31-JUL-2000; 2000US-0221942.		
XX		21-DEC-2000; 2000US-0257600.		
XX		08-JAN-2001; 2001US-0260285.		
XX		(CURA-) CURAGEN CORP.		
XX		Padigaru M, Burgess CE, Majumder K, Mishra VS, Li L;		
XX		Baumgartner JC, Spytek KA, Tchernev VT;		
XX		WPI: 2001-565586/63.		
XX		P-PSDB; AAE09961.		
XX		Novel polypeptides designated as NOVX polypeptides, useful in		
XX		detection, prevention and treatment of e.g. Parkinson's disease and		
XX		Cancer -		
XX		Claim 8; Page 10; 174pp; English.		
XX		The invention relates to novel G-protein coupled-receptor (GPCR) related		
XX		polypeptides and their corresponding nucleotides, referred as NOVX or		
XX		NOV1, NOV2, NOV3, NOV4, NOV5, NOV6, NOV7, NOV8, NOV9 and NOV10. NOVX and		
XX		its Ab are useful for treating or preventing a NOVX-associated disorder,		
XX		such as a disorder related to cell signal processing and metabolic signal		
XX		pathway modulation, e.g. diabetes. Pharmaceutical composition comprising		
XX		NOVX sequence is used to treat or to prevent disorders or syndromes		
XX		including metabolic disturbances associated with obesity, anorexia,		
XX		wasting disorders associated with chronic diseases, infectious diseases		
XX		(particularly infections caused by HIV-1 or HIV-2), cancer-associated		
XX		cachexia, cancer, Parkinson's disease, Alzheimer's disease, immune		
XX		disorders, neurodegenerative diseases and haematopoietic disorders. NOVX		
XX		sequence is also useful for treating developmental diseases, major		
XX		histocompatibility complex (MHC) II and III diseases (immune diseases),		
XX		noninsulin-dependent diabetes mellitus (NIDDM1), bulimia, hypotension,		
XX		acute heart failure, multiple sclerosis, hypertension, osteoporosis,		
XX		Cromb's disease, Albright Hereditary Osteodystrophy, angina pectoris,		
XX		myocardial infarction, asthma, allergies, benign prostatic hypertrophy		
XX		and psychotic and neurological disorders, including anxiety, delirium,		
XX		schizophrenia, manic depression, dementia, severe mental retardation,		
XX		dentaloroballidolusian atrophy (DRLA) hypophosphataemic rickets,		
XX		autosomal dominant (2) acrocallosal syndrome and dyskinesias, such as		
XX		Huntington's disease or Gilles de la Tourette syndrome and/or other		
XX		pathologies. NOVX DNA is used in mapping the chromosomal location of		
XX		NOVX gene and in forensic biology. The present sequence is GPCR NOV1b		
XX		DNA.		
XX		Sequence 943 BP; 221 A; 227 C; 187 G; 308 T; 0 other;		

Query Match	100.0%;	Score 943;	DB 22;	Length 943;
Best Local Similarity	100.0%;	Pred. No. 5.7e-277;		
Matches 943;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Oy	1	tatgatattgggtaaatagacagatatacaagaagattatctctgctggtttctcagatcg	60	
Db	1	tatgatattgggtaaatagacagatatacaagaagattatctctgctggtttctcagatcg	60	
Oy	61	accttggctgtgagtttccaactcctctgtgtgtctcttttgatcttctaacctgtgacaactt	120	
Db	61	accttggctgtgagtttccaactcctctgtgtgtctcttttgatcttctaacctgtgacaactt	120	
Oy	121	tgggaatcggaccattatctctagtgtaacgacctggaaaccaaatctcaatccccatgtga	180	
Db	121	tgggaatcggaccattatctctagtgtaacgacctggaaaccaaatctcaatccccatgtga	180	
Oy	181	ttttcttcttaaccaatctatactcaactcctctggaatctttgtttiaaccaaatgttacaagttccaca	240	
Db	181	ttttcttcttaaccaatctatactcaactcctctggaatctttgtttiaaccaaatgttacaagttccaca	240	
Oy	241	aatgctaagtaaatattatgcaagcatcagaagaagtaatcagttacgtgtgtgtgtgtgccca	300	
Db	241	aatgctaagtaaatattatgcaagcatcagaagaagtaatcagttacgtgtgtgtgtgtgccca	300	
Oy	301	gcttttcacatcttcgggaccttggggccttgaactaataatctctccgggacctgatatccct	360	
Db	301	gcttttcacatcttcgggaccttggggccttgaactaataatctctccgggacctgatatccct	360	
Oy	361	tgatgtgtttgtagcaattctgtcggcctctccatctactcaagtatcatcagtcagacaagaagact	420	
Db	361	tgatgtgtttgtagcaattctgtcggcctctccatctactcaagtatcatcagtcagacaagaagact	420	
Oy	421	ctgacctcaagttgggaagctgtgcatcctctgtgttactgtgttttagtaactaagtggtgttc	480	
Db	421	ctgacctcaagttgggaagctgtgcatcctctgtgttactgtgtgttttagtaactaagtggtgttc	480	
Oy	481	tacctgtgactccagcctgtgcaacctcttggaaacctatgtgaatactcaattctctgtga	540	
Db	481	tacctgtgactccagcctgtgcaacctcttggaaacctatgtgaatactcaattctctgtga	540	
Oy	541	agtcctcgactgtctcaagttatctctgtgtgttgagacaacaagcaaatgtgagtgctgaactat	600	
Db	541	agtcctcgactgtctcaagttatctctgtgtgttgagacaacaagcaaatgtgagtgctgaactat	600	
Oy	601	ccttgtaagtgagcctctccatctataatcccccttgagacatcccttatataatagcttt	660	
Db	601	ccttgtaagtgagcctctccatctataatcccccttgagacatcccttatataatagcttt	660	
Oy	661	tatttccgagagaagatttgaagatacaagctctgtctaagaagtcgacaanaaagcaatttgggac	720	
Db	661	tatttccgagagaagatttgaagatacaagctctgtctaagaagtcgacaanaaagcaatttgggac	720	
Oy	721	atgtgtgtcccaatctaatatgtgtgtctctcttttatagttaagaagcgcgtctgtgtacct	780	
Db	721	atgtgtgtcccaatctaatatgtgtgtctctcttttatagttaagaagcgcgtctgtgtacct	780	
Oy	781	gcaaccacaccttggccagctccaagaagacaagaagaagtgtgtctctctcttaaggaat	840	
Db	781	gcaaccacaccttggccagctccaagaagacaagaagaagtgtgtctctctcttaaggaat	840	
Oy	841	catgtcacccatgtcgaatccccctatatataatacacttgaagaaagaggttaagaaggaag	900	
Db	841	catgtcacccatgtcgaatccccctatatataatacacttgaagaaagaggttaagaaggaag	900	
Oy	901	ctttaaagagtgtgtgtgcaagaagatctctttaatcaagaataa 943		
Db	901	ctttaaagagtgtgtgtgcaagaagatctctttaatcaagaataa 943		

XX AAD17015;
 AC 29-NOV-2001 (first entry)
 DT
 XX G-protein coupled-receptor (GPCR) NOVA DNA.
 DE
 XX G-protein coupled-receptor; GPCR; NOVA-associated disorder; obesity;
 KM anorexia; wasting disorder; infection; cachexia; Parkinson's disease;
 KM Alzheimer's disease; immune disorder; neurodegenerative disease; cancer;
 KM anorectic; haematopoietic disorder; major histocompatibility complex;
 KM MHCII; noninsulin-dependent diabetes mellitus; NIDDM1; bulimia; asthma;
 KM acute heart failure; hypotension; multiple sclerosis; hypertension;
 KM osteoporosis; Crohn's disease; mental retardation; dementia; allergy;
 KM angina pectoris; myocardial infarction; benign prostatic hypertrophy;
 KM psychiatric disorder; neurological disorder; anxiety; schizophrenia;
 KM manic depression; delirium; Albright Hereditary Osteodystrophy; rickets;
 KM dentatorubropallidolysian atrophy; DRPLA; haemostatic; neuroleptic;
 KM autosomal dominant (2) acrocallosal syndrome; dyskinesia; neuroleptic;
 KM Huntington's disease; Gilles de la Tourette syndrome; neuroprotective;
 KM nocturnal; antibacterial; protozoacide; fungicide; cytostatic; NOVA;
 KM vulnery; cardiac; osteopathic; antitumor; antitumor; antitumor;
 KM human immunodeficiency virus; HIV; antidepressant; tranquilliser; ds.
 XX
 OS Unidentified.
 XX
 FH Key Location/Qualifiers
 FH 5'UTR 1..14
 FT /*tag- a
 FT CDS 15..956
 FT /*tag- b
 FT sig-peptide /product- "GPCR NOVA protein"
 FT 15..128
 FT /*tag- c
 FT mat-peptide 129..953
 FT /*tag- d
 FT 957..993
 FT 3'UTR /product- "Mature GPCR NOVA protein"
 FT /*tag- e
 FT
 XX WO20016746-A2.
 XX
 PD 13-SEP-2001.
 XX
 PF 05-MAR-2001; 2001WO-US07116.
 XX
 PR 03-MAR-2000; 2000US-0186606.
 PR 06-MAR-2000; 2000US-0187247.
 PR 06-MAR-2000; 2000US-0187248.
 PR 06-MAR-2000; 2000US-0187249.
 PR 06-MAR-2000; 2000US-0187250.
 PR 06-MAR-2000; 2000US-0187253.
 PR 06-MAR-2000; 2000US-0187295.
 PR 06-MAR-2000; 2000US-0187296.
 PR 07-MAR-2000; 2000US-0187563.
 PR 21-JUL-2000; 2000US-0219854.
 PR 24-JUL-2000; 2000US-0220263.
 PR 31-JUL-2000; 2000US-0221942.
 PR 21-DEC-2000; 2000US-0257600.
 PR 08-JAN-2001; 2001US-0260285.
 XX
 PA (CURA-) CURAGEN CORP.
 XX
 PI Padigaru M, Burgess CE, Majumder K, Mishra VS, Li L;
 PI Baumgartner JC, Spytek KA, Tchervet VT;
 XX WPI: 2001-565586/63.
 DR P-PSDB; AAE09960.
 XX
 XX Novel polypeptides designated as NOVA polypeptides, useful in
 PT detection, prevention and treatment of e.g. Parkinson's disease and
 PT Cancer -

PS Claim 8; Page 8; 174pp; English.
 XX
 CC The invention relates to novel G-protein coupled-receptor (GPCR) related
 CC polypeptides and their corresponding nucleotides, referred as NOVA or
 CC NOVA1, NOVA2, NOVA3, NOVA4, NOVA5, NOVA6, NOVA7, NOVA8, NOVA9 and NOVA10. NOVA and
 CC its Ab are useful for treating or preventing a NOVA-associated disorder,
 CC such as disorder related to cell signal processing and metabolic signal
 CC pathway modulation, e.g. diabetes. Pharmaceutical composition comprising
 CC NOVA sequence is used to treat or to prevent disorders or syndromes
 CC including metabolic disturbances associated with obesity, anorexia,
 CC wasting disorders associated with chronic diseases, infectious diseases
 CC (particularly infections caused by HIV-1 or HIV-2), cancer-associated
 CC cachexia, cancer, Parkinson's disease, Alzheimer's disease, immune
 CC disorders, neurodegenerative diseases and haematopoietic disorders. NOVA
 CC sequence is also useful for treating developmental diseases, major
 CC histocompatibility complex (MHC) II and III diseases (immune diseases),
 CC noninsulin-dependent diabetes mellitus (NIDDM1), bulimia, hypotension,
 CC acute heart failure, multiple sclerosis, hypertension, osteoporosis,
 CC Crohn's disease, Albrit Hereditary Osteodystrophy, angina pectoris,
 CC myocardial infarction, asthma, allergies, benign prostatic hypertrophy
 CC and psychotic and neurological disorders, including anxiety, delirium,
 CC schizophrenia, manic depression, dementia, severe mental retardation,
 CC dentatorubropallidolysian atrophy (DRPLA) hypophosphataemic rickets,
 CC autosomal dominant (2) acrocallosal syndrome and dyskinesias, such as
 CC Huntington's disease or Gilles de la Tourette syndrome and/or other
 CC pathologies. NOVA DNA is used in mapping the chromosomal location of
 CC NOVA gene and in forensic biology. The present sequence is GPCR NOVA
 CC DNA.
 CC
 SQ Sequence 993 BP: 246 A; 233 C; 194 G; 320 T; 0 other:
 XX
 XX
 Query Match 99.7%; Score 940.4; DB 22; Length 993;
 Best Local Similarity 99.9%; Pred. No. 3.6e-276;
 Matches 941; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 2 atgattggggaatgacagatcatatagattatctgcgggtttccagatcga 61
 DB 15 atgattggggaatgacagatcatatagattatctgcgggtttccagatcga 74
 QY 62 ccttgagctggagttccacccctctgtgtctctgtatcttccatgtaacattt 121
 DB 75 ccttgagctggagttccacccctctgtgtctctgtatcttccatgtaacattt 134
 QY 122 ggcacatcgaacattatctagttgcaagcctggaacaaattcacaaccatgat 181
 DB 135 ggcacatcgaacattatctagttgcaagcctggaacaaattcacaaccatgat 194
 QY 182 ttttttaccacatcatcaccctctgtatcttgttaacacacatgtaagtcacaa 241
 DB 195 ttttttaccacatcatcaccctctgtatcttgttaacacacatgtaagtcacaa 254
 QY 242 atgctagtaaatatgagacatcagaagaatgaatcgtgtgtgtgagccag 301
 DB 255 atgctagtaaatatgagacatcagaagaatgaatcgtgtgtgtgagccag 314
 QY 302 ctttccatattctgagccttgaggcctgactgaataattctcctgagcgtcatgtctt 361
 DB 315 ctttccatattctgagccttgaggcctgactgaataattctcctgagcgtcatgtctt 374
 QY 362 gatgtgtttagctatttgcggcctccatcactactgaattatcagacacagaagatc 421
 DB 375 gatgtgtttagctatttgcggcctccatcactactgaattatcagacacagaagatc 434
 QY 422 tgcctcagttgagcagcagcctcctggttactcgttttagtaactcagttgtgtct 481
 DB 435 tgcctcagttgagcagcagcctcctggttactcgttttagtaactcagttgtgtct 494
 QY 482 accctgactcctcagcctgagcactctgagaccatctgtagatcacttctctgtgaa 541
 DB 495 accctgactcctcagcctgagcactctgagaccatctgtagatcacttctctgtgaa 554
 QY 542 gtccctgacactgctcaagttatctgtgtttagagacaacagaatgagcgtgaactatc 601

|||||
Db 555 gtccctcactgtccaagtatctgtgtgagacacacagcaaatgagctgaactatc 614
Qy 602 ctgtcagtagagctcctccatcaataccctcgaactcctcctatatataatgcttt 661
Db 615 ctgtcagtagagctcctccatcaataccctcgaactcctcctatatataatgcttt 674
Qy 662 attgtccgagcagatattgagatacagctcgtcgtgaaggtcgacacaaaagcattgggaca 721
Db 675 attgtccgagcagatattgagatacagctcgtcgtgaaggtcgacacaaaagcattgggaca 734
Qy 722 tgtgtgtcccatcaatattgtgtgtgtcctttttatagtacagcgtcctgtgtacctg 781
Db 735 tgtgtgtcccatcaatattgtgtgtgtcctttttatagtacagcgtcctgtgtacctg 794
Qy 782 caaccaccttcgcgcagctccaaggaaggaagatgtttctctctctatgagatc 841
Db 795 caaccaccttcgcgcagctccaaggaaggaagatgtttctctctctatgagatc 854
Qy 842 attgaccccatgtcgtgaatcccttatatacacacttagaacaagaggtaaaggaagc 901
Db 855 attgaccccatgtcgtgaatcccttatatacacacttagaacaagaggtaaaggaagc 914
Qy 902 tttaaaaggtgtgtgcaagagcttcttataaagaataa 943
Db 915 tttaaaaggtgtgtgcaagagcttcttataaagaataa 956

RESULT 3
AAS08541
ID AAS08541 standard; cDNA; 942 BP.
XX
AC AAS08541;
XX
DT 26-SEP-2001 (first entry)
XX
DE DNA encoding novel human G-protein coupled receptor (NGPCR).
XX
KW Human: novel G-protein coupled receptor; NGPCR; drug discovery;
KW diagnostic; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..942
FT /tag- a
FT /product- "Novel G-protein coupled receptor (NGPCR)"
XX
XX
PN MO200151634-A1.
XX
PD 19-JUL-2001.
XX
PF 05-JAN-2001; 2001MO-US00589.
XX
PR 12-JAN-2000; 2000US-0175764.
XX
PA (LEXI-) LEXICON GENETICS INC.
XX
PI Walke DW, Wiljanowski NL, Turner CA, Friedrich G, Abulin A;
PI Zambrowicz B, Sands AT;
XX
XX WPI; 2001-442145/47.
DR P-PSDB; AAU04689.
XX
XX
PT New human G-protein-coupled receptor and polynucleotides encoding the
PT receptor, useful in identifying, selecting or validating new molecular
PT targets for drug discovery and in diagnostic or prognostic assays -
XX
XX Claim 1; Page 61; 65pp; English.
PS
XX
CC The sequence represents the coding sequence of novel human G-protein
CC coupled receptor (NGPCR). The NGPCR nucleic acid and protein are useful
CC in identifying, selecting or validating novel molecular targets for drug

CC discovery, as well as in diagnostic or prognostic assays. These are also
CC useful in microarrays or other assay formats, for screening collections
CC of genetic material from patients who have a particular medical condition
CC or for identifying mutations associated with a particular disease.
XX
SQ Sequence 942 BP; 222 A; 228 C; 187 G; 305 T; 0 other:

Query Match 99.6%; Score 938.8; DB 22; Length 942;
Best Local Similarity 99.6%; Pred. No. 1.1e-275;
Matches 940; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 atgaattgggtaaatgacagacatcaacagagttattctcgtcgtgttcctcagatcga 61
Db 1 atgaattgggtaaatgacagacatcaacagagttattctcgtcgtgtgttcctcagatcga 60
Qy 62 ccttggtcgtgaattccaccctcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 121
Db 61 ccttggtcgtgaattccaccctcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 120
Qy 122 ggcacatcgcacattatcttagtgcagcctcgtgacacccaactcattaccccatgtat 181
Db 121 ggcacatcgcacattatcttagtgcagcctcgtgacacccaactcattaccccatgtat 180
Qy 182 ttttttcttccaaatcattacacccctcgtgattcctgttaaccacacatgtaagtcacaa 241
Db 181 ttttttcttccaaatcattacacccctcgtgattcctgttaaccacacatgtaagtcacaa 240
Qy 242 atgtcagtaattatgacagacatcaggaagtaacagttatcgtgtgtgtgtgtgtgtgtgt 301
Db 241 atgtcagtaattatgacagacatcaggaagtaacagttatcgtgtgtgtgtgtgtgtgtgt 300
Qy 302 ctcttcataattctcgtgcttgggggtcgtacgtgaatattctcctcgtgcgtcatgtcctt 361
Db 301 ctcttcataattctcgtgcttgggggtcgtacgtgaatattctcctcgtgcgtcatgtcctt 360
Qy 362 gattgtgttttagctatttctcgtgcctcctcattacccagttatcagaccagagatc 421
Db 361 gattgtgttttagctatttctcgtgcctcctcattacccagttatcagaccagagatc 420
Qy 422 tgcctcagcttgagcagctgcatcctcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 481
Db 421 tgcctcagcttgagcagctgcatcctcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 480
Qy 482 accctgacatccagctcagctcagctcgtgacccctatgtatagatcattctctgtgaa 541
Db 481 accctgacatccagctcagctcagctcgtgacccctatgtatagatcattctctgtgaa 540
Qy 542 gtccctcagctgtcacaagtatattctgtgtgtgagacacagcaaatgagctgaactatc 601
Db 541 gtccctcagctgtcacaagtatattctgtgtgtgagacacagcaaatgagctgaactatc 600
Qy 602 ctgtcagtagagctcctccatcaataccctcgaactcctcctatatataatgctttc 661
Db 601 ctgtcagtagagctcctccatcaataccctcgaactcctcctatatataatgctttc 660
Qy 662 attgtccgagcagatattgagatacagctcgtcgtgaaggtcgacacaaaagcattgggaca 721
Db 661 attgtccgagcagatattgagatacagctcgtcgtgaaggtcgacacaaaagcattgggaca 720
Qy 722 tgtgtgtcccatcaatattgtgtgtgtcctttttatagtacagcgtcctgtgtacctg 781
Db 721 tgtgtgtcccatcaatattgtgtgtgtcctttttatagtacagcgtcctgtgtacctg 780
Qy 782 caaccaccttcgcgcagctccaaggaaggaagatgtttctctctctatgagatc 841
Db 781 caaccaccttcgcgcagctccaaggaaggaagatgtttctctctctatgagatc 840
Qy 842 attgaccccatgtcgtgaatcccttatatacacacttagaacaagaggtaaaggaagc 901
Db 841 attgaccccatgtcgtgaatcccttatatacacacttagaacaagaggtaaaggaagc 900
Qy 902 tttaaaaggtgtgtgcaagagcttcttataaagaataa 943

```
|||||
Db 901 tttaaaagttgttgcagaagcttcttataatcaagaataa 942

RESULT 4
ID AAS08542 standard; cDNA; 1488 BP.
AC AAS08542;
XX
XX 26-SEP-2001 (first entry)
XX
XX ORF encoding novel human G-protein coupled receptor (NGPCR).
DE
XX Human; novel G-protein coupled receptor; NGPCR; drug discovery;
KW diagnostic; ss.
XX
XX Homo sapiens.
OS
XX
XX Key Location/Qualifiers
XX CDS 198..1139
XX /*tag= a
XX /product= "Novel G-protein coupled receptor (NGPCR)"
XX
XX NM2000151634-A1.
XX
XX 19-JUL-2001.
XX
XX 05-JAN-2001; 2001MO-US00589.
XX
XX 12-JAN-2000; 2000US-0175764.
XX
XX (LEXI-) LEXICON GENETICS INC.
XX
XX Walke DW, Wilganowski NL, Turner CA, Friedrich G, Abuin A;
XX Zambrowicz B, Sands AT;
XX
XX MPI; 2001-442145/47.
XX
XX P-PSDB; AAU04689.
XX
XX New human G-protein-coupled receptor and polynucleotides encoding the
XX receptor, useful in identifying, selecting or validating new molecular
XX targets for drug discovery and in diagnostic or prognostic assays -
XX
XX PS Disclosure; Page 62-63; 65pp; English.
XX
XX CC The sequence represents the open reading frame (ORF) encoding novel human
XX G-protein coupled receptor (NGPCR). The NGPCR nucleic acid and protein
XX are useful in identifying, selecting or validating novel molecular
XX targets for drug discovery, as well as in diagnostic or prognostic
XX assays. These are also useful in microarrays or other assay formats, for
XX screening collections of genetic material from patients who have a
XX particular medical condition or for identifying mutations associated with
XX a particular disease.
XX
XX SQ Sequence 1488 BP; 421 A; 333 C; 254 G; 477 T; 3 other;

Query Match 99.6%; Score 938.8; DB 22; Length 1488;
Best Local Similarity 99.8%; Pred. No. 1.3e-275;
Matches 940; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
QY 182 ttttttaccatcctcctcctgagcttctgttaaccacacatgacgtcccaaa 241
Db 378 ttttttccacaatcctacacactcctcctgagcttctgttaaccacacatgacgtcccaaa 437
QY 242 atgctaagtaaatatgacagatcagaagaagtaacgtatctgtgtgtgtgagccag 301
Db 438 atgctaagtaaatatgacagatcagaagaagtaacgtatctgtgtgtgtgagccag 497
QY 302 ctcttcataattctgtgctctgtgggagctactgaatactctcctgcgcgcgtcatctctt 361
Db 498 ctcttcataattctgtgctctgtgggagctactgaatactctcctgcgcgcgtcatctctt 557
QY 362 gattggtttgtagctattctgtgcgcctccatcactcactcactcactcactcactcactc 421
Db 558 gatagttttagtattctgtgcgcctccatcactcactcactcactcactcactcactcactc 617
QY 422 tgacctcagttgagctcgtcactcctcctggttactggtttttagtaactaagtgtgtgtct 481
Db 618 tgacctcagttgagctcgtcactcctcctggttactggtttttagtaactaagtgtgtgtct 677
QY 482 accctgactctccagctcgtcactcctcgtgagccctcgtgatactcactcctcgtgaa 541
Db 678 accctgactctccagctcgtcactcctcgtgagccctcgtgatactcactcctcgtgaa 737
QY 542 gtccctgactcgtcactcgttctgtgtgtgagacaacagcaaatgaggtgaaactatc 601
Db 738 gtccctgactcgtcactcgttctgtgtgtgagacaacagcaaatgaggtgaaactatc 797
QY 602 ctgtcagttgagctctccatcactaataaccctgagactcactccttatacatagctttt 661
Db 798 ctgtcagttgagctctccatcactaataaccctgagactcactccttatacatagctttt 857
QY 662 atgtccagagcagattgagagatacagctcgtgaaggtcgacaacaaagcatttgggaca 721
Db 858 atgtccagagcagattgagagatacagctcgtgaaggtcgacaacaaagcatttgggaca 917
QY 722 tgtgttcccatcctaattgtgtgtctcttttatagtacagcgtctcgtgtactcg 781
Db 918 tgtgttcccatcctaattgtgtgtctcttttatagtacagcgtctcgtgtactcg 977
QY 782 caaccacttgcgcagctcctcaagagcaagaagaagtgttctcttctataggaatc 841
Db 978 caaccacttgcgcagctcctcaagagcaagaagaagtgttctcttctataggaatc 1037
QY 842 atgcaccatgctgaatcccttatatatatacacttgaagcaagaagtaagaagc 901
Db 1038 atgcaccatgctgaatcccttatatatatacacttgaagcaagaagtaagaagc 1097
QY 902 tttaaaagttgttgcagaagctcttctaatacaagaataa 943
Db 1098 tttaaaagttgttgcagaagctcttctaatacaagaataa 1139

RESULT 5
AAH32018
ID AAH32018 standard; DNA; 939 BP.
XX
XX AAH32018;
AC
XX 30-JUL-2001 (first entry)
XX
XX Human olfactory receptor polynucleotide, SEQ ID NO: 591.
DE
XX Human; olfactory receptor; OR: primary scent determination;
KW secondary scent determination; polypeptide library; odour receptor;
KW scent profile; scent fingerprint; scent representation; ds.
XX
XX Homo sapiens.
OS
XX
XX NM2000127158-A2.
XX
XX PD 19-APR-2001.
```

XX 06-OCT-2000; 2000WQ-US27582.
 PF 08-OCT-1999; 99US-0158615.
 PR 24-FEB-2000; 2000US-0184809.
 XX (DIGI-) DIGISCENTS.
 PA (YEDA) YEDA RES & DEV CO LTD.
 XX Ballenson J, Smith D, Lancet D, Glusman G, Fuchs T, Yanai I;
 PI WPI; 2001-290713/30.
 DR New polynucleotides which encode polypeptides involved in olfactory
 PT sensation for identifying olfactory agonists and antagonists -
 XX
 PS Claim 8; Page 417; 1857pp; English.
 XX
 CC The present sequence is one of a number of isolated polynucleotides
 CC which encode polypeptides involved in olfactory sensation. The
 CC polynucleotides can be used in screening for olfactory agonists and
 CC antagonists. The methods allow for the determination of primary
 CC scents and the identification of the odour receptors used to detect
 CC these primary scents. The methods also enable determination of
 CC secondary scents and the identification of combinations of odour
 CC receptors that are involved in detecting such secondary scents.
 CC This enables the construction of a scent representation (also called
 CC a scent fingerprint or scent profile), which may be used to re-create
 CC and edit scents. Libraries of olfactory receptors are useful for
 CC determining the interaction pattern of a composition with the receptors,
 CC and can be used for determining differences in the olfactory faculties
 CC of different individuals.
 CC
 SQ Sequence 939 BP; 220 A; 227 C; 187 G; 305 T; 0 other;

Query Match 99.4%; Score 937.4; DB 22; Length 939;
 Best Local Similarity 99.9%; Pred. NO. 2.9e-275;
 Matches 938; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 atgaaatgggtaaaatgacagacatatacagagattatctctgctgggttcacagatga 61
 Db 1 atgattgggtaaaatgacagacatatacagagattatctctgctgggttcacagatga 60
 QY 62 ccttgctggaggttcac 121
 Db 61 ccttgctggaggttcac 120
 QY 122 ggcacatgacacattatctagctgacgctggacacacacacacacacacacacacat 181
 Db 121 ggcacatgacacattatctagctgacgctggacacacacacacacacacacacacat 180
 QY 182 tttttcttcccaatcaccac 241
 Db 181 tttttcttcccaatcaccac 240
 QY 242 atgctagtaaatatcagac 301
 Db 241 atgctagtaaatatcagac 300
 QY 302 cttttcatattctggcctgggggctgactgaaatatacttcccgccgcgtatgctctt 361
 Db 301 cttttcatattctggcctgggggctgactgaaatatacttcccgccgcgtatgctctt 360
 QY 362 gattggtctgagattctgctggcctctccattacacagattatcagacacacagagact 421
 Db 361 gattggtctgagattctgctggcctctccattacacagattatcagacacacagagact 420
 QY 422 tgcctcagattgagcagctgac 481
 Db 421 tgcctcagattgagcagctgac 480
 QY 482 accctgactctccagcgcac 541

Db 481 accctgactctccagcgcac 540
 QY 542 gtccctcagctcgaagtattctctgtgtgagacacacacacacacacacacacacacacac 601
 Db 541 gtccctcagctcgaagtattctctgtgtgagacacacacacacacacacacacacacacac 600
 QY 602 ctgtcagtgagctcttccatcataacacacacacacacacacacacacacacacacacacac 661
 Db 601 ctgtcagtgagctcttccatcataacacacacacacacacacacacacacacacacacacac 660
 QY 662 atgtccgagcagattgagatagacagctgctgtaaggtcgacacacacacacacacacac 721
 Db 661 atgtccgagcagattgagatagacagctgctgtaaggtcgacacacacacacacacacac 720
 QY 722 ttgtgtcccatcgaattgtgtgtctctttttatagtaacgctgctgttactgt 781
 Db 721 ttgtgtcccatcgaattgtgtgtctctttttatagtaacgctgctgttactgt 780
 QY 782 caaccacctcgcacagctcccaagacacacacacacacacacacacacacacacacacacac 841
 Db 781 caaccacctcgcacagctcccaagacacacacacacacacacacacacacacacacacacac 840
 QY 842 atgcacacacatgctgaattcccttatatacacttgtagaacacacacacacacacacac 901
 Db 841 atgcacacacatgctgaattcccttatatacacttgtagaacacacacacacacacacac 900
 QY 902 tttaaaggtgtgtgcaagagctcttcttaacaagaa 940
 Db 901 tttaaaggtgtgtgcaagagctcttcttaacaagaa 939

RESULT 6
 AAD17017
 ID AAD17017 standard; DNA; 956 BP.
 XX
 AC AAD17017;
 XX
 DT 29-NOV-2001 (first entry)
 XX
 DE G-protein coupled-receptor (GPCR) NOVIC DNA.
 XX
 KW G-protein coupled-receptor; GPCR; NOVIC-associated disorder; obesity;
 KW anorexia; wasting disorder; infection; cachexia; Parkinson's disease;
 KW Alzheimer's disease; immune disorder; neurodegenerative disease; cancer;
 KW anorectic; haematopoietic disorder; major histocompatibility complex;
 KW MHCII; noninsulin-dependent diabetes mellitus; NIDDM; bulimia; asthma;
 KW acute heart failure; hypotension; multiple sclerosis; hypertension;
 KW osteoporosis; Crohn's disease; mental retardation; dementia; allergy;
 KW angina pectoris; myocardial infarction; benign prostatic hypertrophy;
 KW psychotic disorder; neurological disorder; anxiety; schizophrenia;
 KW manic depression; delirium; Alzright Hereditary Osteodystrophy; rickets;
 KW dentatorubropallidolusian atrophy; DRPLA; haemostatic; anticonvulsant;
 KW autosomal dominant (2) acrocraliosal syndrome; dyskinesia; neuroleptic;
 KW Huntington's disease; Gilles de la Tourette syndrome; neuroprotective;
 KW nootropic; antibacterial; protozoacide; fungicide; cytostatic; NOVIC;
 KW vulnerary; cardiant; osteopathic; antianginal; antidiacer; vincloide;
 KW human immunodeficiency virus; HIV; antidepressant; tranquilliser; ds.
 XX
 OS Unidentified.
 XX
 FH Key Location/Qualifiers
 FT 1..6
 FT 5'UTR
 FT CDS
 FT 7..948
 FT /*tag- a
 FT /*tag- b
 FT /*tag- c
 FT /*tag- d
 FT mat_peptide
 FT 121..945
 FT /*tag- g
 FT /*tag- h
 FT 3'UTR
 FT 946..956
 FT "Mature GPCR NOVIC protein"

```
FT      /*tag- e
XX
XX      WO200166746-A2.
XX
XX      13-SEP-2001.
XX
XX
XX      05-MAR-2001; 2001WO-US07116.
XX
XX      03-MAR-2000; 2000US-0186606.
XX      06-MAR-2000; 2000US-0187247.
XX      06-MAR-2000; 2000US-0187248.
XX      06-MAR-2000; 2000US-0187249.
XX      06-MAR-2000; 2000US-0187250.
XX      06-MAR-2000; 2000US-0187253.
XX      06-MAR-2000; 2000US-0187295.
XX      06-MAR-2000; 2000US-0187296.
XX      07-MAR-2000; 2000US-0187563.
XX      21-JUL-2000; 2000US-0219854.
XX      24-JUL-2000; 2000US-0220263.
XX      31-JUL-2000; 2000US-0221942.
XX      21-DEC-2000; 2000US-0257600.
XX      08-JAN-2001; 2001US-0260285.
XX
XX      (CURA-) CURAGEN CORP.
XX
XX      Padigaru M, Burgess CE, Majumder K, Mishra VS, Li L;
XX      Baumgartner JC, Spytek KA, Tchernev VT;
XX
XX      WPI: 2001-565586/63.
XX      P-PSDB; AAE09962.
XX
XX      Novel polypeptides designated as NOVX polypeptides, useful in
XX      detection, prevention and treatment of e.g. Parkinson's disease and
XX      Cancer -
XX
XX      Claim 8; Page 12; 17app; English.
XX
XX      The invention relates to novel G-protein coupled-receptor (GPCR) related
XX      polypeptides and their corresponding nucleotides, referred as NOVX or
XX      NOV1, NOV2, NOV3, NOV4, NOV5, NOV6, NOV7, NOV8, NOV9 and NOV10. NOVX and
XX      its Ab are useful for treating or preventing a NOVX-associated disorder,
XX      such as disorder related to cell signal processing and metabolic signal
XX      pathway modulation, e.g. diabetes. Pharmaceutical composition comprising
XX      NOVX sequence is used to treat or to prevent disorders or syndromes
XX      including metabolic disturbances associated with obesity, anorexia,
XX      wasting disorders associated with chronic diseases, infectious diseases
XX      (particularly infections caused by HIV-1 or HIV-2), cancer-associated
XX      cachexia, cancer, Parkinson's disease, Alzheimer's disease, immune
XX      disorders, neurodegenerative diseases and haematopoietic disorders. NOVX
XX      sequence is also useful for treating developmental diseases, major
XX      histocompatibility complex (MHC) II and III diseases (immune diseases),
XX      noninsulin-dependent diabetes mellitus (NIDDM), bulimia, hypotension,
XX      acute heart failure, multiple sclerosis, hypertension, osteoporosis,
XX      Crohn's disease, Alibright Hereditary Osteodystrophy, angina pectoris,
XX      myocardial infarction, asthma, allergies, benign prostatic hypertrophy,
XX      CC and psychotic and neurological disorders, including anxiety, delirium,
XX      schizophrenia, manic depression, dementia, severe mental retardation,
XX      denturobiopallidolysian atrophy (DRPLA) hypophosphataemic rickets,
XX      autosomal dominant (2) acrocallosal syndrome and dyskinisias, such as
XX      Huntington's disease or Gilles de la Tourette syndrome and/or other
XX      pathologies. NOVX DNA is used in mapping the chromosomal location of
XX      NOVX gene and in forensic biology. The present sequence is GPCR NOV1C
XX      DNA.
XX
XX      Sequence 956 BP; 227 A; 228 C; 190 G; 311 T; 0 other;
XX
XX
XX      Query Match      99.4%: Score 937.2; DB 22; Length 956;
XX      Best Local Similarity 99.7%: Pred. NO. 3.4e-275;
XX      Matches 939; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
XX
XX      2 atgaattgggaaatgacagcatcatcagaagattatctctgctgggttctcagatcga 61
XX      |||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db      7 atgaattgggaaatgacagcatcatcagaagattatctctgctgggttctcagatcga 66
QY      62 ccttgctggaattccactcctctgtgtctcttgaattcttaccatgtgaccattt 121
|||
Db      67 ccttgctggaattccactcctctgtgtgtctcttgaattcttaccatgtgaccattt 126
QY      122 ggaactctgacattatctagtgctcagcctggaacacaaacttaccaccaatgat 181
|||
Db      127 ggaactctgacattatctagtgctcagcctggaacacaaacttaccaccaatgat 186
QY      182 ttttttaccactatcatcctcctgtgattctgttaccacacatgttaagttccacaa 241
187 ttttcttaccactatcatcctcctgtgattctgttaccacacatgttaagttccacaa 246
QY      242 atgctagtaaatatgatcagatcaggaagaatgaatcgtatcgtgtgtgtagccag 301
247 atgctagtaaatatgatcagatcaggaagaatgaatcgtatcgtgtgtgtagccag 306
Db      302 ctcttcataattctgacctggtgggtactgaataatcttctcgtgacctatgtccttt 361
307 ctcttcataattctgacctggtgggtactgaataatcttctcgtgacctatgtccttt 366
QY      362 gatgtgttgtagctattgtgacctccttcatactaccatgaattacatgaccagaagatc 421
367 gatagtttgtagctattgtgacctccttcatactaccatgaattacatgaccagaagatc 426
QY      422 tgcctcagattgtagctatcctggtgtactgtgttttagtaactcagtggtgtgtct 481
427 tgcctcagattgtagctatcctggtgtactgtgttttagtaactcagtggtgtgtct 486
Db      482 accctgactctcagctgtagcactctgtgacccctatgtatagatcaacttctctgtgaa 541
487 accctgactctcagctgtagcactctgtgacccctatgtatagatcaacttctctgtgaa 546
QY      542 gtccctcagctcctcagaattatctgtgtgtgagacaagaagaatgaggtggaactatc 601
547 gtccctcagctcctcagaattatctgtgtgtgagacaagaagaatgaggtggaactatc 606
QY      602 ctgtcagtgagctcttcacatcatataccctgtagcactcactcatatcatatgctttt 661
607 ctgtcagtgagctcttcacatcatataccctgtagcactcactcatatcatatgctttt 666
QY      662 atgttcgagagatgagatgatacagtcctgtgaaggtgcaaaaaagcgttgggaca 721
667 atgttcgagagatgagatgatacagtcctgtgaaggtgcaaaaaagcgttgggaca 726
Db      722 tgtgttccatctaatgtgtgtctcttttatagtagcagcgtctgtgtactgt 781
727 tgtgttccatctaatgtgtgtctcttttatagtagcagcgtctgtgtactgt 786
QY      782 caacacacttgcacagctcacaagacccaagaagaatggtttctctctatgaaatc 841
787 caacacacttgcacagctcacaagacccaagaagaatggtttctctctatgaaatc 846
QY      842 attgacacacttgcacagctcacaagacccaagaagaatggtttctctctatgaaatc 901
847 attgacacacttgcacagctcacaagacccaagaagaatggtttctctctatgaaatc 906
QY      902 tttaaaaggtgtgtgcaagaagcttctttaaacaagaataa 943
907 tttaaaaggtgtgtgcaagaagcttctttaaacaagaataa 948
Db
RESULT 7
AAH31669
ID AAH31669 standard; DNA; 1071 BP.
XX
XX AAH31669;
AC
XX 30-JUL-2001 (first entry)
DT
XX
XX Human olfactory receptor polynucleotide, SEQ ID NO: 242.
XX
```

KM Human: olfactory receptor; OR; primary scent determination;
 KW secondary scent determination; polypeptide library; odour receptor;
 XX scent profile; scent fingerprint; scent representation; ds.
 OS Homo sapiens.
 XX MO200127158-A2.
 XX 19-APR-2001.
 XX 06-OCT-2000; 2000MO-US27582.
 XX 08-OCT-1999; 99US-0158615.
 XX 24-FEB-2000; 2000US-0184809.
 PA (DIGI-) DIGISCENTS.
 PA (YEDA) YEDA RES & DEV CO LTD.
 PI Bellenson J, Smith D, Lancet D, Glusman G, Fuchs T, Yanai I;
 DR WPI; 2001-290713/30.
 XX New polynucleotides which encode polypeptides involved in olfactory
 PT sensation for identifying olfactory agonists and antagonists -
 XX Claim 8; Page 273-274; 1857pp; English.
 XX The present sequence is one of a number of isolated polynucleotides
 CC which encode polypeptides involved in olfactory sensation. The
 CC polynucleotides can be used in screening for olfactory agonists and
 CC antagonists. The methods allow for the determination of primary
 CC scents and the identification of the odour receptors used to detect
 CC these primary scents. The methods also enable determination of
 CC secondary scents and the identification of combinations of odour
 CC receptors that are involved in detecting such secondary scents.
 CC This enables the construction of a scent representation (also called
 CC a scent fingerprint or scent profile), which may be used to re-create
 CC and edit scents. Libraries of olfactory receptors are useful for
 CC determining the interaction pattern of a composition with the receptors,
 CC and can be used for determining differences in the olfactory faculties
 CC of different individuals.
 XX
 XX Sequence 1071 BP; 277 A; 248 C; 201 G; 345 T; 0 other;
 SQ

Query Match 77.0%; Score 726.2; DB 22; Length 1071;
 Best Local Similarity 86.3%; Pred. No. 7e-211;
 Matches 814; Conservative 0; Mismatches 128; Indels 1; Gaps 1;

QY 2 atgaattgggttaaatgacagcatatcagaggttattctgtggtttctcagatcga 61
 DB 1 atgaattgggttaaatgagatgtccacagaggttattctgtggtttctcagatcga 60
 QY 62 ccttggttgaggttccacctgtgtctgttctgttcttcttcttcttcttcttctt 121
 DB 61 ccattggttgaggttccacctgtgtctgttctgttcttcttcttcttcttcttctt 120
 QY 122 ggcacatcgacatattctagttgacagcgttgacacacaaacttaccacccatgtat 181
 DB 121 ggcacatcgacatattctgttgcacatgttgattcacaaccacacccatgtat 180
 QY 182 tttttcttccaactatcactctgtgattctgttacaacacatgtacagttccacaa 241
 DB 181 tttttcttccaactatcactctgtgattctgttacaacacatgtacagttccacaa 240
 QY 242 atgctagtaattatcagcatcagatcaggaagaatcagttatcgtgtgtgtgtccag 301
 DB 241 atgctagtaattatcagcatcagatcaggaagaatcagttatcgtgtgtgtgtccag 300
 QY 302 ctcttcattctctgcttggtggtggtcactgaatatctctcctggtcgtcatgtcctt 361
 DB 301 ctcttcattctctgcttggtggtggtcactgaatatctctcctggtcgtcatgtcctt 360

QY 362 gatgtgtttgtagctatttgcggcctcccatcactcactgaattatcagtcacagagctc 421
 DB 361 gatagttttgtagctatttgcggcctcccatcactcactgaattatcagtcacagagctc 420
 QY 422 tgcctcagtttggagcttgatccttggtgttacttggttttgaactaagttgtgtct 481
 DB 421 tgcctcagtttggagcttgatccttggtgttacttggttttgaactaagttgtgtct 480
 QY 482 accctgactctcagctgcacactctgtgacacccatgtgatatcacttctctgtgaa 541
 DB 481 accctgactctcagctgcacactctgtgacacacaaagtggttcaacttctctgtgaa 540
 QY 542 gtccctgacactgacagatctctgtgtgagacacaaagtggttcaacttctctgtgaa 601
 DB 541 gtccctgacactgacagatctctgtgtgagacacaaagtggttcaacttctctgtgaa 600
 QY 602 ctgtcaggtgagctcttcacatctaacccctgacacatcacttatacgtcttt 661
 DB 601 ctgtcaggtgagctcttcacatctaacccctgacacatcacttatacgtcttt 660
 QY 662 atgtccgagcagatttggagatagacagctctgtgaaagtcgacaaagaatttggaca 721
 DB 661 atgtccgagcagatttggagatagacagctctgtgaaagtcgacaaagaatttggaca 720
 QY 722 tgtgttccatctaatgtgtgtctcttttataagcagcgtctgtgtactctg 781
 DB 721 tgtgttccatctaatgtgtgtctcttttataagcagcgtctgtgtactctg 780
 QY 782 caaccactctgcacacgtccaaagacaaagaaagtgttctctctcactgaatc 841
 DB 781 caaccactctgcacacgtccaaagacaaagaaagtgttctctctcactgaatc 840
 QY 842 atgcacacatgctgaatcccttataatatacacttgagaaagagtgtaaaagagc 901
 DB 841 atgcacacatgctgaatcccttataatatacacttgagaaagagtgtaaaagagc 900
 QY 902 tttaaaagttgtgtgc-aagagtccttctaatacagaataa 943
 DB 901 tttaaaagttgtgtgc-aagagtccttctaatacagaataa 943

RESULT 8
 ID AAH32250 standard; DNA; 1071 BP.
 XX
 XX AAH32250;
 AC
 XX
 DT 30-JUL-2001 (first entry)
 XX
 XX Human olfactory receptor polynucleotide, SEQ ID NO: 823.
 DE
 XX Human: olfactory receptor; OR; primary scent determination;
 KW secondary scent determination; polypeptide library; odour receptor;
 KW scent profile; scent fingerprint; scent representation; ds.
 OS Homo sapiens.
 XX MO200127158-A2.
 XX 19-APR-2001.
 XX 06-OCT-2000; 2000MO-US27582.
 XX 08-OCT-1999; 99US-0158615.
 XX 24-FEB-2000; 2000US-0184809.
 PA (DIGI-) DIGISCENTS.
 PA (YEDA) YEDA RES & DEV CO LTD.
 PI Bellenson J, Smith D, Lancet D, Glusman G, Fuchs T, Yanai I;
 DR WPI; 2001-290713/30.
 XX

PT New polynucleotides which encode polypeptides involved in olfactory
 PT sensation for identifying olfactory agonists and antagonists -
 XX
 XX Claim 8; Page 508-509; 1857pp; English.
 XX
 CC The present sequence is one of a number of isolated polynucleotides
 CC which encode polypeptides involved in olfactory sensation. The
 CC polynucleotides can be used in screening for olfactory agonists and
 CC antagonists. The methods allow for the determination of primary
 CC scents and the identification of the odour receptors used to detect
 CC these primary scents. The methods also enable determination of
 CC secondary scents and the identification of combinations of odour
 CC receptors that are involved in detecting such secondary scents.
 CC This enables the construction of a scent representation (also called
 CC a scent fingerprint or scent profile), which may be used to re-create
 CC and edit scents. Libraries of olfactory receptors are useful for
 CC determining the interaction pattern of a composition with the receptors,
 CC and can be used for determining differences in the olfactory faculties
 CC of different individuals.
 XX
 XX Sequence 1071 BP; 277 A; 248 C; 201 G; 345 T; 0 other;

Query Match 77.0%; Score 726.2; DB 22; Length 1071;
 Best Local Similarity 86.3%; Pred. No 7e-211;
 Matches 814; Conservative 0; Mismatches 128; Indels 1; Gaps 1;

QY 2 atgaattggttaaatgacacatatacagagatttattctgctgggtttctcagatca 61
 DB 1 atgaattggttaaatgacacatatacagagatttattctgctgggtttctcagatca 60
 QY 62 ccttggtgagattcaccctctgctgcttcttctgatttcttcaactgagacatttt 121
 DB 61 ccaagtgtagatcaccctctgctgcttcttctgatttcttcaactgagacatttt 120
 QY 122 ggaacatggaacattatctagttgacgctgagacacaaacttaccacccattat 181
 DB 121 ggaacatggaacattatctggttgcacatggtgatttcaaacccacacccatgac 180
 QY 182 ttttttctacacatatacactctcctgattctgttaccacacatgtagtccacaa 241
 DB 181 ttttttctacacatatacactctcctgattctgttaccacacatgtagtccacaa 240
 QY 242 atgtagtaaatatgacagatcagagaagtaacagttatctgctgctggtgacccag 301
 DB 241 atgctggttaaacacatatacagagaagtaacagttatctgctgctggtgacccag 300
 QY 302 ctttcaatttctgctgctggtggtgacacatatacctcctggcgctgacgtcctt 361
 DB 301 ctttcaatttctgctgctggtggtgacacagatgtcttctcctggcgctgacgtcctt 360
 QY 362 gattggttgaagctatttgcgctcctccatcattcaagttatcagacagagatc 421
 DB 361 gatagtttgagctatttgcgctcctccatcattcaagttatcagacagagatc 420
 QY 422 tgcctcaagttgagacgtcgtcctggttactggttttagtaactcaagtggtgtct 481
 DB 421 tgcctcaagttgagacgtcgtcctggttactggttttagtaactcaagttgacgtc 480
 QY 482 accctgactcagcagctcgtgacacccatgtgtagagtaactcttctcgtgaa 541
 DB 481 accctgactcagcagctcgtgacacccatgtgtagagtaactcttctcgtgaa 540
 QY 542 gtccctgacgtcacaagtattctgtgtgagacaacagcaaatgaggtgacatttc 601
 DB 541 gtccctgacgtcacaagtattctgtgtgagacaacagcaaatgaggtgacatttc 600
 QY 602 ctgtgagtgagctcctccatataacccctgacactcatccttatacatatgctttt 661
 DB 601 ctgtgagtgagctcctccatataacccctgacactcatccttatacatatgctttt 660
 QY 662 atgtccagagagattgagatagctctctgaaagtgtagcaaaaagccttggtgaca 721
 DB 662 atgtccagagagattgagatagctctctgaaagtgtagcaaaaagccttggtgaca 721

DB 661 atgtccagagagattgagatagctctctgaaagtgtagcaaaaagccttggtgaca 720
 QY 722 tgtgttccatcctaattgtgtgtctcttctttatagtagcaagcgtctctgttactcg 781
 DB 721 tgtgttccatcctaattgtgtgtgtctcttctttatagtagcaagcgtctctgttactcg 780
 QY 782 caaccacttcgcccagctcctcacaagaggaagtgtttcttctctatgagtc 841
 DB 781 caaccacttcacccagctcctcacaagaggaagtgtttcttctctctgtgagtc 840
 QY 842 atgtccacatgctgaaatcccttatatacacttagtaggaacagaggtgaaagagc 901
 DB 841 atgtccacatgctgaaatcccttatatacacttagtaggaacagaggtgaaagagc 900
 QY 902 tttaaaagttgtgtgc-aagagcttcttaatacaagaataa 943
 DB 901 tttaaaagttgtgtgcaagaagcttcttaatacaagaataa 943

RESULT 9
 AAH32231 standard; DNA; 648 BP.
 ID
 XX
 XX AAH32231;
 AC
 XX
 XX 30-JUL-2001 (first entry)
 DT
 XX
 XX Human olfactory receptor polynucleotide, SEQ ID NO: 804.
 DE
 XX
 XX Human: olfactory receptor; OR: primary scent determination;
 KW secondary scent determination; polypeptide library; odour receptor;
 KW scent profile; scent fingerprint; scent representation; ds.
 OS
 XX Homo sapiens.
 XX
 XX WO200127158-A2.
 PN
 XX
 XX 19-APR-2001.
 PD
 XX
 XX 06-OCT-2000; 2000MO-US27582.
 PF
 XX
 XX 08-OCT-1999; 99US-0158615.
 PR 24-FEB-2000; 2000US-0184809.
 XX
 XX (DIGI-) DIGISCENTS.
 PA
 PA (YEDA) YEDA RES & DEV CO LTD.
 PI
 XX
 XX Bellenson J, Smith D, Lancet D, Glusman G, Fuchs T, Yanai I;
 DR WPL; 2001-290713/30.
 XX
 XX
 XX New polynucleotides which encode polypeptides involved in olfactory
 PT sensation for identifying olfactory agonists and antagonists -
 XX
 XX Claim 8; Page 502; 1857pp; English.
 PS

XX The present sequence is one of a number of isolated polynucleotides
 CC which encode polypeptides involved in olfactory sensation. The
 CC polynucleotides can be used in screening for olfactory agonists and
 CC antagonists. The methods allow for the determination of primary
 CC scents and the identification of the odour receptors used to detect
 CC secondary scents and the identification of combinations of odour
 CC receptors that are involved in detecting such secondary scents.
 CC This enables the construction of a scent representation (also called
 CC a scent fingerprint or scent profile), which may be used to re-create
 CC and edit scents. Libraries of olfactory receptors are useful for
 CC determining the interaction pattern of a composition with the receptors,
 CC and can be used for determining differences in the olfactory faculties
 CC of different individuals.
 XX
 XX Sequence 648 BP; 144 A; 166 C; 132 G; 206 T; 0 other;

Query Match 68.0%; Score 641.6; DB 22; Length 648;
Best Local Similarity 99.4%; Pred. No. 3,4e-185;
Matches 644; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 203 ctcttgatcttcttaccacatgtacagtcaccacaagtctagtaaatatgcagc 262
DB 1 ctcttgatcttcttaccacatgtacagtcaccacaagtctagtaaatatgcagc 60
QY 263 atcaggaagaatcagttacagtcgctgtgtacagtcaccagcttccattcttcgctg 322
DB 61 atcaggaagaatcagttacagtcgctgtgtacagtcaccagcttccattcttcgctg 120
QY 323 ggggactagtaatactctctcgcgcgtacgtcccttgattgttctgtacatttgt 382
DB 121 ggggactagtaatactctctcgcgcgtacgtcccttgattgttctgtacatttgt 180
QY 383 cggcctcctcattcattatcatatgcacagagactctgcctcagttggagctgca 442
DB 181 cggcctcctcattcattatcatatgcacagagactctgcctcagttggagctgca 240
QY 443 tccctgggttactgttttaagtaactcagttgtgtgtctaccctcagcttccagctgca 502
DB 241 tccaggttactcgttcttaagtaactcagttgtgtgtctaccctcagcttccagctgca 300
QY 503 ctctgtgaccctatgtgatagatcattctctgtgaagtcctcgtcagctgcaagtt 562
DB 301 ctctgtgaccctatgtgatagatcattctctgtgaagtcctcgtcagctgcaagtt 360
QY 563 tctgtgttggagaagaagaatgaagtggaactatctctgtcagtgagcttccat 622
DB 361 tctgtgttggagaagaagaatgaagtggaactatctctgtcagtgagcttccat 420
QY 623 ctataccctcgtacacatcctctatcatatgcttcttattgtccagagcagttag 682
DB 421 ctataccctcgtacacatcctctatcatatgcttcttattgtccagagcagttag 480
QY 683 ataacgctcgtcgaagtcgacaaaagcatttggagacatgtgttcccatcaattgtg 742
DB 481 ataacgctcgtcgaagtcgacaaaagcatttggagacatgtgttcccatcaattgtg 540
QY 743 gtgtcctttttttatagcagcgtctctgtgtacactgcaacaccttgcagagctc 802
DB 541 gtgtcctttttttatagcagcgtctctgtgtacactgcaacaccttgcagagctc 600
QY 803 aaggaccaaaggaagatggttctctctctctcattgaatcattgaccc 850
DB 601 aaggaccaaaggaagatggttctctctctctcattgaatcattgaccc 648

RESULT 10
AAC77475
ID AAC77475 standard; CDNA; 1442 BP.
XX AAC77475:
AC AAC77475:
DT 08-FEB-2001 (first entry)
DE Human ORFX ORF3030 polynucleotide sequence SEQ ID NO:6059.
XX
XX
XX Human: open reading frame; ORFX: detection; cytostatic; hepatotropic;
KM vulnereary; antipsoiatric; antiparkinsonian; nootropic; neuroprotective;
KM anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;
KM immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
KM hypotensive; dermatological; immunosuppressive; antiinflammatory;
KM antiviral; antibacterial; antifungal; antirheumatic; antihypertensive;
KM antineoplastic; cancer; proliferative disorder; hypertension;
KM neurodegenerative disorder; osteoarthritis; graft vs host disease;
KM cardiovacular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
KM cholesterol ester storage; systemic lupus erythematosus; infection;
KM severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
KM allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
KM bone damage; cartilage damage; antiinflammatory disease; coagulation;

KW thrombosis; contraceptive; ss.
XX
XX Homo sapiens.
XX
XX WO200058473-A2.
XX
XX
XX 05-OCT-2000.
XX
XX 31-MAR-2000; 2000WO-US08621.
XX
XX 31-MAR-1999; 99US-0127607.
XX 02-APR-1999; 99US-0127636.
XX 05-APR-1999; 99US-0127728.
XX 30-MAR-2000; 2000US-0540763.
XX
XX (CURA-) CURAGEN CORP.
XX
XX Shinkets RA, Leach M;
XX
XX WPI; 2000-602362/57.
XX P-PSDB; ABA3266.
XX
XX Novel nucleic acids and peptides derived from open reading frame X,
PT useful for treating e.g. cancers, proliferative disorders,
PT neurodegenerative disorders and cardiovascular disease -
XX
XX Claim 5; Page 5245-5246; 5507Pp; English.

CC AAC74446 to AAC77606 encode the proteins given in ABA40237 to ABA43397,
CC which represent the human ORFX open reading frames 1 to 3161. The ORFX
CC sequences have activities such as: cytostatic; hepatotropic; vulnereary;
CC antipsoiatric; antiparkinsonian; nootropic; neuroprotective;
CC osteopathic; anticonvulsant; antiarthritic; immunosuppressant;
CC immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;
CC antidiabetic; hypotensive; dermatological; immunosuppressive;
CC antiinflammatory; antibacterial; antiviral; antifungal; antirheumatic;
CC antihypertensive; antineoplastic. The sequences can be used for determining
CC the presence of or predisposition to, or preventing or treating
CC pathological conditions associated with an ORFX-associated disorder. The
CC nucleic acids can be used to express ORFX proteins in gene therapy
CC vectors. The proteins and nucleic acids may be used to treat cancers,
CC proliferative disorders, neurodegenerative disorders, osteoarthritis,
CC graft vs host disease, cardiovacular disease, diabetes mellitus,
CC hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,
CC allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
CC nocturnal haemoglobinuria, antiinflammatory disease; to enhance
CC coagulation; to inhibit thrombosis; and as a contraceptive.
XX
XX Sequence 1442 BP; 378 A; 319 C; 268 G; 477 T; 0 other;

Query Match 56.6%; Score 534; DB 21; Length 1442;
Best Local Similarity 72.9%; Pred. No. 3e-152;
Matches 687; Conservative 0; Mismatches 255; Indels 0; Gaps 0;

QY 2 atgaattggtaaatgacagcatcatcaggaatttctctgctgggttcttcagatcga 61
DB 301 atgaattggtaaatgacagcatcatcaggaatttctctgctgggttcttcagatcga 360
QY 62 ccttgctcgtgagttcaccctcctgtgtcttcttgattcttaccatcgtgacacattc 121
DB 361 gcttggttacaagaagcccttctgtgtctgttatacatcacaacacacattt 420
QY 122 gccaatcgtacattatcttagtctcagcgtcgtgacacaaacttaccacacattat 181
DB 421 gccaatcgtcaccatcagatggtgtcattctgtgattcccaactcattccattat 480
QY 182 ttcttctacatcatcactcctcgtgattctgttaccacacatgtacagtcaccaa 241
DB 481 ttcttctacatcatcctcctcattcttagatctcgtcattaccacacacattccat 540

OY 242 atgctagtaattatgacgacatcagaagaatcaatcagttatcgtgctgtgtagccag 301
 DB 541 atgctgtaaatatcgttgcacaacaaagaccatctatgttgcgtgtgtgtagccac 600
 OY 302 cttctcatattctgcgcttgggagctactgaataatcttcctgcgcgcgtatccctt 361
 DB 601 ctacatactctccgcctcagtagtgcacagaagtgctcctcctgcgctgtatgtctt 660
 OY 362 gattgtttgtagctatttgcgctctccatcactactactgaatcagtcacagagactc 421
 DB 661 gacagatatggtgtgttgtagacccctccactatgtagcatcagtaattatgtgttc 720
 OY 422 tgcctcagttgagagctgcactcctgtgttactgtgttttagtaactcagttgtgtct 481
 DB 721 tgcctagagagagagagctcctcctcctcctcctcctcctcctcctcctcctcctc 780
 OY 482 accctgactcctcagcgcgcctcctcctcctcctcctcctcctcctcctcctcctcctc 541
 DB 781 tccctgactcctcactacagcgcgcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 840
 OY 542 gtccctgactgctcagagttatcttgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 601
 DB 841 gtgcctgactcctcactcagagttgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 900
 OY 602 ctgtgtagtgcctcctcactcactcactcactcactcactcactcactcactcactcact 661
 DB 901 tctcttgtagcactcactcctcactcactcactcactcactcactcactcactcactcact 960
 OY 662 atgtccgagcagatgtgagagatcagctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 721
 DB 961 atagctcaagagagatgt 1020
 OY 722 tgt 781
 DB 1021 tgt 1080
 OY 782 caaccctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 841
 DB 1081 caaccctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 1140
 OY 842 atgtcaccatgt 901
 DB 1141 atcaccatcagtgtagt 1200
 OY 902 tttaaaggt 943
 DB 1201 ttcaagagctgt 1242
 RESULT 11
 AAH32304
 ID AAH32304 standard; DNA; 939 BP.
 XX
 AC AAH32304;
 XX
 DT 30-JUL-2001 (first entry)
 XX
 DE Human olfactory receptor polynucleotide, SEQ ID NO: 877.
 XX
 KM Human: olfactory receptor; OR: primary scent determination;
 KM secondary scent determination; polypeptide library; odour receptor;
 KM scent profile; scent fingerprint; scent representation; ds.
 XX
 OS Homo sapiens.
 XX
 PN MO200127158-A2.
 XX
 PD 19-APR-2001.
 XX
 PF 06-OCT-2000; 2000MO-US27582.
 XX
 PR 08-OCT-1999; 99US-0158615.
 PR 24-FEB-2000; 2000US-0184809.

XX
 PA (DIGT-) DIGISCENTS.
 PA (YEDA) YEDA RES & DEV CO LTD.
 XX
 PI Bellenson J, Smith D, Lancet D, Glusman G, Fuchs T, Yanai I;
 XX
 DR WPL: 2001-290713/30.
 XX
 PT New polynucleotides which encode polypeptides involved in olfactory
 PT sensation for identifying olfactory agonists and antagonists -
 XX
 PS Claim 8; Page 530; 1857pp; English.
 XX
 CC The present sequence is one of a number of isolated polynucleotides
 CC which encode polypeptides involved in olfactory sensation. The
 CC polynucleotides can be used in screening for olfactory agonists and
 CC antagonists. The methods allow for the determination of primary
 CC scents and the identification of the odour receptors used to detect
 CC these primary scents. The methods also enable determination of
 CC secondary scents and the identification of combinations of odour
 CC receptors that are involved in detecting such secondary scents.
 CC This enables the construction of a scent representation (also called
 CC a scent fingerprint or scent profile), which may be used to re-create
 CC and edit scents. Libraries of olfactory receptors are useful for
 CC determining the interaction pattern of a composition with the receptors,
 CC and can be used for determining differences in the olfactory faculties
 CC of different individuals.
 XX
 SQ Sequence 939 BP; 232 A; 228 C; 185 G; 294 T; 0 other;
 Query Match 56.3%; Score 531; DB 22; Length 939;
 Best Local Similarity 72.8%; Pred. No. 2e-151;
 Matches 684; Conservative 0; Mismatches 255; Indels 0; Gaps 0;
 OY 2 atgaattgggaatgaagcagatcagagatgtatcgtcgtgtgtgtgtgtgtgtgtgtgtgt 61
 DB 1 atgaattgggaatgaagcagatcagagatgtatcgtcgtgtgtgtgtgtgtgtgtgtgtgt 60
 OY 62 ccttgcgtgaggttccactcctgt 121
 DB 61 gcttgcgtcacaagcgccttctgt 120
 OY 122 ggcacatcagcatatctagtgtagcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 181
 DB 121 ggcacatcagcatatctagtgtagcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 180
 OY 182 ttttttaccatatactactcctcgt 241
 DB 181 ttttttaccatatactactcctcgt 240
 OY 242 atgctagtaaatatgacgacatcaggaagtaacagttacgtgtgtgtgtgtgtgtgtgtgt 301
 DB 241 atgtgtgtaaatatgt 300
 OY 302 ctttcatattctgcgcttgggagctactgaatacttctcctgcgcgcgtatcctt 361
 DB 301 ctcatcactcctgcgcgcgtactgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 360
 OY 362 gattgtttgtagctatttgcgctcctcctcactcactcactcactcactcactcactcactc 421
 DB 361 gacagatatggt 420
 OY 422 tgcctcagttgagagctgcactcctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 481
 DB 421 tgcctagagagagagagctcctcctcctcctcctcctcctcctcctcctcctcctcctc 480
 OY 482 accctgactcctcagcgcgcctcctcctcctcctcctcctcctcctcctcctcctcctc 541
 DB 481 tccctgactcctcactacagcgcgcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 540
 OY 542 gtccctgactgctcagagttatcttgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 601

KW cerebroprotective; G-coupled receptor; cell proliferative disease;
KW lymphoma; leukemia; breast cancer; cirrhosis; neurological disorder;
KW stroke; Alzheimer's disease; multiple sclerosis; mental retardation;
KW cardiovascular disease; atherosclerosis; angina pectoris; indigestion;
KW congestive heart failure; gastrointestinal disorder; dysphagia; AIDS;
KW gastritis; autoimmune disorder; inflammatory disorder; Crohn's disease;
KW systemic lupus erythematosus; metabolic disorder; diabetes; obesity;
KW viral infection; herpesvirus; parvovirus;
KW acquired immune deficiency syndrome; ss.
XX
OS Homo sapiens.
XX
PN WO200190359-A2.
XX
PD 29-NOV-2001.
XX
PF 22-MAY-2001: 2001MO-US16833.
XX
PR 22-MAY-2000: 2000US-206222P.
PR 25-MAY-2000: 2000US-207476P.
PR 02-JUN-2000: 2000US-208834P.
PR 02-JUN-2000: 2000US-208861P.
PR 07-JUN-2000: 2000US-209868P.
XX
PA (INCYTE) INCYTE GENOMICS INC.
XX
PI Patterson C, Tribouley CM, Yao MG, Griffin JA, Thornton M, Lu Y;
PI Kallick DA, Gandhi AR, Au-Young J;
XX
DR MPI: 2002-106199/14.
XX P-PSDB: AA080497.
XX
PT New G-protein coupled receptors useful for treating or preventing cell
PT proliferative (e.g. leukemia), neurological (e.g. stroke),
PT cardiovascular or autoimmune/inflammatory disorders -
XX
XX
PS Claim 5: Page 140; 148bp; English.
XX
XX The invention relates to a novel human G-coupled receptor (I). (I) and
CC its corresponding polynucleotides are useful for diagnosing, treating or
CC preventing cell proliferative diseases (e.g. lymphoma, leukemia, breast
CC cancer or cirrhosis), neurological disorders (e.g. stroke, Alzheimer's
CC disease, multiple sclerosis or mental retardation), cardiovascular
CC diseases (e.g. atherosclerosis, angina pectoris or congestive heart
CC failure), gastrointestinal disorders (e.g. dysphagia, indigestion or
CC gastritis), autoimmune/inflammatory disorders (e.g. AIDS, Crohn's disease
CC or systemic lupus erythematosus) or metabolic disorders (e.g. diabetes
CC or obesity), or viral infections (e.g. infection by herpesvirus or
CC parvovirus), ABK16615-ABK16637 represent novel human G-coupled
CC receptor coding sequences of the invention.
XX
SQ Sequence 939 BP; 223 A; 226 C; 189 G; 301 T; 0 other;

Query Match 42.6%; Score 402; DB 24; Length 939;
Best Local Similarity 65.0%; Pred. No. 4..1e-112;
Matches 594; Conservative 0; Mismatches 320; Indels 0; Gaps 0;

QY 2 atgaattggaatgacagcatcatcaggaatttatcttgctgggttctcagacga 61
DB 1 atggatcagaagaatggaagttcttcaatcgattatccctacgggttcttcgacag 60
QY 62 ccttgctggaagttcaccacctgtgtgtctctctgattcttctacactgtgacatctt 121
DB 61 cctcagctgtagctagccctctgtgtgtctcttctgtctctctacatctcactctgtc 120
QY 122 ggcattcgcacattatctagtgtcaagcctggaaccaactcatalaccccatgtat 181
DB 121 gggacaacaacacatcatgattatctcacttgaccacacatctcactcctatgtat 180
QY 182 ttttcttaccatcatcactcctggtatcttgttaaccacacatgtaagtcacaa 241
DB 181 ttttcttccacaaagcttttggatctgttaccacacacgcgcatgttccacag 240

QY 242 atgtatgaattatcagcatcaggaagtaatcagttatctgtgctgtgtacccag 301
DB 241 ctctgtttaaattccacaggagacagacaataatccatctcctatgtgtgtgtgtccag 300
QY 302 ctcttcatattctgacctgtggggactagtaataatctctcctgacctatgctctt 361
DB 301 ctgacatctctctcagcttggagatctacagaatgctctctcttagagatgtatgtatt 360
QY 362 gatgtgtttagtatttctggtcctctccattaccatcagttatcagacacagagctc 421
DB 361 gaccgcatgcaagctgttgcagagccctccacacacagtaagtcagacacctgtctg 420
QY 422 tgcctcagttgagcagctacccctgtggttactgtgtttagtaagcagtggtgtct 481
DB 421 tatgtgtgtatggtcttacttacttcaatgtgtatgtgttttgcaactccctatgcaagc 480
QY 482 accctgactctcagctcagctcagctgtgacccctatgtatagatcactctctgtgaa 541
DB 481 gtgtcatcttgcctttaacacatttggagaagaataatagaacaccttctgtgag 540
QY 542 gtccctcagctcagctcaggtatcttgtgttgagacacacagcaatgagcgtgacatctc 601
DB 541 gtccctcagctcagctcagctcagctgtgtgacacactctatgtaatctgacatcttc 600
QY 602 ctgtcagttgagctcttccatctaatccctgacactacatccttatcatatgtcttt 661
DB 601 ttgtcagttgattatcttcttctgttactcgttgcattatcatatctctcatagtcag 660
QY 662 atgtccgagcagatgtgagatcagatctgtctgtaagtcgacaaagcattgtggaca 721
DB 661 atgtcagggcagatcagatgataagttagcaacagggcagagaagatgtttggaca 720
QY 722 tgtgtcccatcctaattgtgtgtctcttttttatgtaacagcgtctgtgtactgt 781
DB 721 tgtgtcccatcctaattgtgtgtctcttttttatgtaacagcgtctgtgtactgt 780
QY 782 caaccactctgcagctcagctcaggaaggaacaggaagatgttctctctatgtgaatc 841
DB 781 cagcccggaacaactactctcagatcagatcaggaagtcattctctctctacacatc 840
QY 842 atgtcaccatgtcgaatccctctatatataacttggagaacagggatgaaggaagc 901
DB 841 attacacacatgtacacacccctctatatataacttggagaacaggaatgtgaaaggaca 900
QY 902 tttaaaagttgtgt 915
DB 901 cttaagaagttgtct 914

RESULT 15
AAH32063
ID AAH32063 standard; DNA: 1002 BP.
XX
XX AAH32063:
AC
XX
DT 30-JUL-2001 (first entry)
XX
DE Human olfactory receptor polynucleotide, SEQ ID NO: 636.
XX
XX Human olfactory receptor; OR: primary scent determination;
KW secondary scent determination; polypeptide library; odour receptor;
KW scent profile; scent fingerprint; scent representation; ds.
XX
OS Homo sapiens.
XX
PN WO200127158-A2.
XX
PD 19-APR-2001.
XX
PF 06-OCT-2000: 2000MO-US27582.
XX
PR 08-OCT-1999: 9905-0158615.

PR 24-FEB-2000: 2000US-0184809.
 XX (DIGI-) DIGISCENTS.
 PA (YEDA) YEDA RES & DEV CO LTD.
 XX
 PI Bellenson J, Smith D, Lancelot D, Glusman G, Fuchs T, Yanai I;
 XX WPI; 2001-290713/30.
 XX
 PT New polynucleotides which encode polypeptides involved in olfactory
 PT sensation for identifying olfactory agonists and antagonists -
 XX
 PS Claim 8; Page 435-436; 1857pp; English.
 XX
 XX The present sequence is one of a number of isolated polynucleotides
 CC which encode polypeptides involved in olfactory sensation. The
 CC polynucleotides can be used in screening for olfactory agonists and
 CC antagonists. The methods allow for the determination of primary
 CC scents and the identification of the odour receptors used to detect
 CC these primary scents. The methods also enable determination of
 CC secondary scents and the identification of combinations of odour
 CC receptors that are involved in detecting such secondary scents.
 CC This enables the construction of a scent representation (also called
 CC a scent fingerprint or scent profile), which may be used to re-create
 CC and edit scents. Libraries of olfactory receptors are useful for
 CC determining the interaction pattern of a composition with the receptors,
 CC and can be used for determining differences in the olfactory faculties
 CC of different individuals.
 XX
 SO Sequence 1002 BP; 210 A; 266 C; 217 G; 309 T; 0 other;

Query Match 39.4%; Score 371.6; DB 22; Length 1002;
 Best Local Similarity 62.5%; Pred. No. 7.8e-103;
 Matches 581; Conservative 0; Mismatches 349; Indels 0; Gaps 0;

QY 14 aatgacagctatcacagaggttatctctggtgttcacagatcgacctgtgag 73
 DB 73 aatgagagcaacctagagaggttccattcttaggtttcttgcattacccacgttacag 132
 QY 74 ttccacctctgtgtctctcttgcattcttaccctgtgacacattcttggaactgacc 133
 DB 133 aaggtctattgttgccatattgactgtatcttactaatttttgggaataacacc 192
 QY 134 attattcagtgacgcctgacacccaattacacccacatgtaatttttcttacc 193
 DB 193 atcattctgttctcgttcgaacccaagcttcacatagtcgatgtatttcttcttct 252
 QY 194 aatctacactcctcgtgattcttctgttaacacacatgtaagtcacacaatgctagtaaat 253
 DB 253 catctcctcctcgtcgtcgtcttaccagcagtgtaattccacagctcctcgtgtaaac 312
 QY 254 ttatgcagcatcaggaagaatcagtaacgtacgtgtagccagctttcatattt 313
 DB 313 ctgtggaacacacgaataactatgcctatgtagtgctgttgcattacacttaacactcc 372
 QY 314 ctggccttgggggactgaatatctctccctggcgcacatgctccttgattgttgta 373
 DB 373 catgccttgatccactgagtgctccctccgctgtagtctcgtgtaacgctatctg 432
 QY 374 gctattgtgcgcctcctcattacatgatacatgacacagagactcgtccagttg 433
 DB 433 gctgtcgcgcctcctcctcattacacgttcttaagtatatcactcttgatgaccttg 492
 QY 434 gcaagctgacactcgtgtagtctttagtaacacagtgtagtcttaccctactctc 493
 DB 493 gcatctatgcatagggcctcagtgatgacacacccctgtaacagtcacactcaccctg 552
 QY 494 cagctgcacactctgtgacccctatgtgatacatcttctctgtgaagtcctcgtcagctg 553
 DB 553 cagctgcactctgtgtggcagtcgcaagtgtgatacttcatctgcagagtcctcgtctc 612
 QY 554 ctcaagtattctgtgtgtgacacagcaaatgtaggctgaactatctcctgttcaagtga 613

DB 613 atcaagctggtctgtgtggtgacacacgtttaaagagctgagcttctgtgctagatc 672
 QY 614 ctctccatctaaatccccgacacactcatccttatcatatgcttttatgtctcggagca 673
 DB 673 ctcttccttatagtgctcgtcctcatcactatcctgtctccctcctgctacatgtcccaagca 732
 QY 674 gttatgagatcacagctcgtcgtgaagtcgacaaaagacatttggacatgtgttccat 733
 DB 733 ggttgaggttaagtcagctacacaggaagaacgacatttcggacctgtctccac 792
 QY 734 ctattgtgtgtctctttttatagtaacagcgtctcctgtgtacccgacacctg 793
 DB 793 ctgacagtggttcaaccacttttaagacacatcactcactgtaactcgcagccagcaag 852
 QY 794 cccagctccaaagcaaggaagatgttctctcttctatgtaacatgacacccatg 853
 DB 853 agtagatcccgaggaacggcaggttcttctctctccttcaacatggtgtaacccgcatg 912
 QY 854 ctgaatcccttatatatatacacattaggaacaagaggttaaggaagctttaaaggttg 913
 DB 913 cttaaccttatattatcacttgagatacaagaggtgaaagggcattaaagaagttc 972
 QY 914 gttgcaagagcttcttaatacaagaataa 943
 DB 973 ctgacaaagctcgtggagtaaatatttta 1002

Search completed: August 25, 2002, 19:02:13
 Job time: 4201 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 25, 2002, 17:41:07 : Search time 77.06 Seconds
(without alignments)
3005.871 Million cell updates/sec

Title: US-09-800-321A-3
Perfect score: 943
Sequence: 1 tatgaattggtaataagca.....tcttctaataagaataaa 943

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 38353 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA: *
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq: *
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq: *
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq: *
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq: *
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq: *
6: /cgn2_6/ptodata/2/ina/backfile1.seq: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	215.2	22.8	1290	2 US-08-827-291A-1	Sequence 1, Appl
2	213.2	22.6	966	3 US-08-748-506-7	Sequence 7, Appl
3	212.6	22.5	966	3 US-08-748-506-8	Sequence 8, Appl
4	201.6	21.4	900	4 US-09-085-371-5	Sequence 5, Appl
5	199.2	21.1	1713	2 US-08-467-948A-1	Sequence 1, Appl
6	199.2	21.1	1713	3 US-08-467-947A-1	Sequence 1, Appl
7	198.8	21.1	966	3 US-08-748-506-5	Sequence 5, Appl
8	192	20.4	966	3 US-08-748-506-6	Sequence 6, Appl
9	185.2	19.6	984	3 US-08-748-506-9	Sequence 9, Appl
10	120	12.7	963	4 US-09-439-313-526	Sequence 526, App
11	116.8	12.4	1474	1 US-08-465-980-1	Sequence 1, Appl
12	116.8	12.4	1474	1 US-09-053-303-1	Sequence 1, Appl
13	116.8	12.4	1474	5 PCT-US95-07093-1	Sequence 2, Appl
14	96	10.2	1828	3 US-08-988-876-2	Sequence 1, Appl
15	90.2	9.6	1320	1 US-08-599-252-84	Sequence 84, Appl
16	90.2	9.6	1320	1 US-08-436-074-57	Sequence 57, Appl
17	90.2	9.6	1320	5 PCT-US96-06352-84	Sequence 84, Appl
18	90.2	9.6	1320	5 PCT-US96-06583-84	Sequence 84, Appl
19	59	6.3	7218	1 US-08-232-463-14	Sequence 14, Appl
20	41.6	4.4	1137	3 US-09-082-088-1	Sequence 1, Appl
21	41.6	4.4	1137	4 US-09-546-117-1	Sequence 1, Appl
22	41.4	4.4	1260	1 US-08-599-252-83	Sequence 83, Appl
23	41.4	4.4	1260	1 US-08-436-074-56	Sequence 56, Appl
24	41.4	4.4	1260	5 PCT-US96-06352-83	Sequence 83, Appl
25	39.8	4.2	2012	3 US-08-706-281A-7	Sequence 7, Appl
26	39.8	4.2	2012	3 US-08-706-281A-7	Sequence 7, Appl
27	39.8	4.2	2012	4 US-09-097-231-7	Sequence 7, Appl

C	28	37	3.9	251	4	US-09-439-313-352	Sequence 352, App
	29	36.6	3.9	1015	1	US-08-671-525B-3	Sequence 3, Appl
	30	36.6	3.9	1015	1	US-08-672-109B-3	Sequence 3, Appl
	31	36.6	3.9	1015	1	US-08-842-045-3	Sequence 3, Appl
	32	36.6	3.9	1015	2	US-08-842-238-3	Sequence 3, Appl
	33	36.6	3.9	1015	3	US-08-629-335B-3	Sequence 3, Appl
	34	36.6	3.9	2012	1	US-07-866-560-3	Sequence 3, Appl
	35	36.6	3.9	2012	1	US-08-077-673-3	Sequence 3, Appl
	36	36.6	3.9	2012	1	US-08-478-992-3	Sequence 3, Appl
	37	36.6	3.9	2012	3	US-09-105-298-3	Sequence 3, Appl
C	38	34	3.6	289	4	US-09-007-005-17	Sequence 17, Appl
	39	34	3.6	289	4	US-09-244-796-17	Sequence 17, Appl
C	40	31.8	3.4	1101	3	US-08-945-056-5	Sequence 5, Appl
C	41	30.8	3.3	753	4	US-08-998-416-574	Sequence 5, Appl
C	42	30.4	3.2	603	3	US-08-609-334-10	Sequence 10, Appl
C	43	30.4	3.2	603	3	US-08-609-334-15	Sequence 15, Appl
	44	30.4	3.2	1092	4	US-09-077-675A-15	Sequence 15, Appl
	45	30.4	3.2	3129	4	US-09-077-675A-14	Sequence 14, Appl

ALIGNMENTS

RESULT 1
US-08-827-291A-1
Sequence 1, Application US/08827291A
Patent No. 5874243
GENERAL INFORMATION:
APPLICANT: Macina, Roberto
TITLE OF INVENTION: NOVEL OLRCC15 RECEPTOR
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESSES:
ADDRESSEE: SmithKline Beecham Corporation
STREET: 709 Swedeland Road
CITY: King of Prussia
STATE: PA
COUNTRY: PA
ZIP: 19406
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/827,291A
FILING DATE: 28-MAR-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: King, William T
REGISTRATION NUMBER: 30,954
REFERENCE/DOCKET NUMBER: GP50001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-5015
TELEFAX: 610-270-5090
TELEX:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1290 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-827-291A-1
Query Match 22.8%; Score 215.2; DB 2; Length 1290;
Best Local Similarity 52.1%; Pred. No. 1.9e-57;
Matches 481; Conservative 0; Mismatches 443; Indels 0; Gaps 0;

QY 2 atgaattgggttaaatgacagcatatacaaggagttatctctgcygglltctcagatcga 61
 Db 296 ATGGCAAGGGGAGATTCGACCTTCACCTCCGACTTCCTGCGGATTCGATCAATC 355
 QY 62 ccttgagctgagttccacccctctgctgctctgattcttctacactgagccattt 121
 Db 356 AGCCCCACCCACACCTTCCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 415
 QY 122 ggcacatgcacacattatctagtcacgcccgcgagacacacacacacacacacat 181
 Db 416 GGAACACCTGTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 475
 QY 182 ttttcttaccacatctacacacacacacacacacacacacacacacacacacac 241
 Db 476 CTCTCTCTGACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 535
 QY 242 atgctagtaattatgacagcatagagaagtaacagttatctgctgctgctgctg 301
 Db 536 ATGGCTTCACACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 595
 QY 302 ctctcatattctgagccttgagggcctacacacacacacacacacacacacacac 361
 Db 596 ATTTCTCTGATACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 655
 QY 362 gattggttgcagctatctgctgctgctgctgctgctgctgctgctgctgctg 421
 Db 656 GACCGCTACACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 715
 QY 422 tgcctcagttgagcctgacacacacacacacacacacacacacacacacacac 481
 Db 716 TGTGGACTTATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 775
 QY 482 accctgac 541
 Db 776 GTACCCACATTTCT 835
 QY 542 gtccctgac 601
 Db 836 TTACCTTCCT 895
 QY 602 ctgtcagttgagcctctcactcactcactcactcactcactcactcactcactc 661
 Db 896 ATTGGCTCTATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 955
 QY 662 attgcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 721
 Db 956 GTTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1015
 QY 722 tgggttcac 781
 Db 1016 TGTTCCT 1075
 QY 782 caac 841
 Db 1076 CAGCCACATCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1135
 QY 842 attgcac 901
 Db 1136 CT 1195
 QY 902 ttttaaaagttggtgagagtc 925
 Db 1196 TTCTATGAAGATCTCTGAGAAAGGC 1219

US-08-748-506-7
 ; Sequence 7, Application US/08748506
 ; Patent No. 613707
 ; GENERAL INFORMATION:
 ; APPLICANT: Ronnelt et al.
 ; TITLE OF INVENTION: NOVEL SPERM RECEPTORS
 ; NUMBER OF SEQUENCES: 31

; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Leydig, Volt & Mayer, Ltd.
 ; STREET: Two Prudential Plaza, Suite 4900
 ; CITY: Chicago
 ; STATE: IL
 ; COUNTRY: US
 ; ZIP: 60601-6780
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/748,506
 ; FILING DATE: 08-NOV-1996
 ; CLASSIFICATION: 435
 ; PRIORITY APPLICATION DATA:
 ; APPLICATION NUMBER: US 60/033,751
 ; FILING DATE: 09-NOV-1995
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; REFERENCE/DOCKET NUMBER: 74940
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 312-616-5600
 ; TELEFAX: 312-616-5700
 ; INFORMATION FOR SEO ID NO: 7:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 966 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: Linear
 ; MOLECULE TYPE: DNA (genomic)
 ; US-08-748-506-7

Query Match 22.6%; Score 213.2; DB 3; Length 966;
 Best Local Similarity 55.4%; Pred. No. 6.3e-57;
 Matches 413; Conservative 0; Mismatches 333; Indels 0; Gaps 0;
 QY 164 cttaac 223
 Db 181 CTACACACCCCACTGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 240
 QY 224 acatgac 283
 Db 241 TGTCT 300
 QY 284 cgtgagctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 343
 Db 301 GAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
 QY 344 ctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 403
 Db 361 TTGGCAGCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
 QY 404 atcatgac 463
 Db 421 CGAATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
 QY 464 aactcagctgagctgagcagcagcagcagcagcagcagcagcagcagcagcag 523
 Db 481 GTAGGCTGGGACACACATTTTATTTTCTCTGTAACCTTCTGTGACCTGTGAATA 540
 QY 524 gatacattctctgagcagcagcagcagcagcagcagcagcagcagcagcagcagc 583
 Db 541 GACCACTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
 QY 584 aatgagcctgac 643
 Db 601 AATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
 QY 644 ctatatacatgcttttatatgctgagcagcagcagcagcagcagcagcagcagcagc 703

[illegible]

:	APPLICANT:	SUTTON III, GRANGER G.
:	APPLICANT:	ROSEN, CRAIG A.
:	TITLE OF INVENTION:	Polynucleotides Encoding Human G-Protein
:	TITLE OF INVENTION:	Coupled Receptor GPR2
:	NUMBER OF SEQUENCES:	30
:	CORRESPONDENCE ADDRESS:	
:	ADDRESSEE:	STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
:	STREET:	1100 NEW YORK AVE., NW, SUITE 600
:	CITY:	WASHINGTON
:	STATE:	DC
:	COUNTRY:	USA
:	ZIP:	20005
:	COMPUTER READABLE FORM:	
:	MEDIUM TYPE:	FLOPPY DISK
:	COMPUTER:	IBM PC COMPATIBLE
:	OPERATING SYSTEM:	PC-DOS/MS-DOS
:	SOFTWARE:	PATENTIN RELEASE #1.0, VERSION #1.30
:	CURRENT APPLICATION DATA:	
:	APPLICATION NUMBER:	US/08/467,948A
:	FILING DATE:	06-JUN-1995
:	CLASSIFICATION:	435
:	PRIOR APPLICATION DATA:	
:	APPLICATION NUMBER:	PCT/US95/04079
:	FILING DATE:	30-MAR-1995
:	ATTORNEY/AGENT INFORMATION:	
:	NAME:	STEEFE, ERIC K.
:	REGISTRATION NUMBER:	36,688
:	REFERENCE/DOCKET NUMBER:	1488..1140003/EKS/KLM
:	TELECOMMUNICATION INFORMATION:	
:	TELEPHONE:	202-371-2600
:	TELEFAX:	202-371-2540
:	INFORMATION FOR SEQ ID NO:	1:
:	SEQUENCE CHARACTERISTICS:	
:	LENGTH:	1713 base pairs
:	TYPE:	nucleic acid
:	STRANDEDNESS:	both
:	TOPOLOGY:	both
:	MOLECULE TYPE:	CDNA
:	FEATURE:	
:	NAME/KEY:	CDS
:	LOCATION:	116..1003
:	US-08-467-948A-1	
:	Query Match	21.1%; Score 199.2; DB 2; Length 1713;
:	Best Local Similarity	52.3%; Pred. No.2.le-52;
:	Matches 441: Conservative	0; Mismatches 403; Indels 0; Gaps 0;
Oy	13	aataacagcatcacacaggatctattctcgtcggtttctcaagtcgaccttgctga 72
Dd	136	AAATGAGCAATGGTCACAGAAGTTCTCCTACTGTGGATTTTCCGCGGCCAACGAATTCA 195
Oy	73	gttccactccttgtgtctctcttgattcttaacactgtgcacatccttggcaatcga 132
Dd	196	GATGCTCCTCTTTGGGCTCTCTCCCTGTTCTCATGCTTCACCCGTCGCGGGAATGGGAC 255
Oy	133	cattctctagtgtacagccctggagacaacaattcatacccccatgatatctttcttac 192
Dd	256	CATCTTGCGGCATCATCTCACTGGAAGCTCCAGACTCCACACCOCCTGTAATCTTCTCTGC 315
Oy	193	caatctatactccttgaatcttgtttacacaacatgatcagtlccccaagaatgtagtaa 252
Dd	316	ACACCTGGCCGCHGCACATGCGCTATGCGTCGCAACAACAGTGCACAGATGCTGTGAA 375
Oy	253	ttaatgcagcatcagaaagtaatacagttatctgtgtcgtgtgtagcccagctttcata 312
Dd	376	CCTCCTGCATCCAGCAAGCCCATCTCCTTGGCTTGTCATGACATGACATGACTTCTCTT 435
Oy	313	tctgaccttgaggagctaatgataatctctcctcggcgctcaatgtccttgaattgtgt 372
Dd	436	TTTAGGATTTCACATATCTGAATGCCCTCTGTTGCTGCTGAATGCTCTACAGATGGTACGT 495
Oy	373	agctattgttcggcctctccattactcagtatcatcagcacagagactctgcctcaglt 432

Db	496	GGCCATCTGGCACCCCTCTCCGATATTTTCATCATCATCACTGACCTGGAAAGTGTGCATCACTCT	555
Qy	433	ggcagctgcgaatccctgggttacttggttttaagtaactcaagtggtgtgtctctacccctgactc	492
Db	556	GGGCATCATCTTCCTGGACATGTGGCTCTCCCTCTGGCTATGCTCCATGTGGGCTCATCTCT	615
Qy	493	ccaagctgcacactcgttaccacctatgtagatagatcaatcttctcgttgaagtcctgact	552
Db	616	AAGACTGCGCTTTGTGTGGGCGCTGCTGAATCAACCACTTCTTCTGTGAATTCCTGTCTGT	675
Qy	553	gctcaagttatcttctgtgttttagacaacagcaaatagagctgaactatcctgttcagtga	612
Db	676	CCTGAGCTGGCCCTGTCTGTATACCTGGCTCAACCAAGTGGTGATCTTTAAACCTTGCA	735
Qy	613	gctcttcacatcaataccctcgacacccatccatctatactaatgattttagtcttcgagc	672
Db	736	GTTCATCTGCTGGGACCACTCTGCTGGTGTCTCTACTCAACATCTCGGGGG	795
Qy	673	agatttagagatacagctcgtgaagctgcacaaaagacatttggagacatgttgcctcca	732
Db	796	CATCTGTAGATTCAGTCTGGGGAGGGCCGAGAAAGCCTTTCACACCTGCTCTCCCA	855
Qy	733	tctaatgtgtgtctctttttatagtagacgcgtctctgtgtacctgaacacacttc	792
Db	856	CCTCTGCGTATGTGGACTCTTCTTTGGASACCGCATCGTACATGTGACCTGCCCCCTAAGTC	915
Qy	793	gccccagctccaaaggacccaagaagaatggtctctctctctatagaaatcatgcaacct	852
Db	916	CCGCGATCCTGAGAGGACGACGAGAAAGTCTTTTCTTATTATTTACAGTCTTTTCAACCC	975
Qy	853	gctg 856	
Db	976	GATG 979	
RESULT 6			
US-08-467-947A-1			
: Sequence 1, Application US/08467947A			
: Patent No. 6090575			
: GENERAL INFORMATION:			
: APPLICANT: LI, YI			
: APPLICANT: CAO, LITANG			
: APPLICANT: NI, JIAN			
: APPLICANT: GENTZ, REINER			
: APPLICANT: BULL, CAROL J.			
: APPLICANT: SUTTON III, GRANGER G.			
: APPLICANT: ROSEN, CRAIG A.			
: TITLE OF INVENTION: Polynucleotides Encoding Human G-Protein			
: TITLE OF INVENTION: Coupled Receptor GPR1			
: NUMBER OF SEQUENCES: 30			
: CORRESPONDENCE ADDRESS:			
: ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.			
: STREET: 1100 NEW YORK AVE., NW, SUITE 600			
: CITY: WASHINGTON			
: STATE: DC			
: COUNTRY: USA			
: ZIP: 20005			
: COMPUTER READABLE FORM:			
: MEDIUM TYPE: FLOPPY DISK			
: COMPUTER: IBM PC COMPATIBLE			
: OPERATING SYSTEM: PC-DOS/MS-DOS			
: SOFTWARE: PATENTIN RELEASE #1.0, VERSION #1.30			
: CURRENT APPLICATION DATA:			
: APPLICATION NUMBER: US/08/467,947A			
: FILING DATE: 06-JUN-1995			
: CLASSIFICATION: 435			
: PRIOR APPLICATION DATA:			
: APPLICATION NUMBER: PCT/US95/04079			
: FILING DATE: 30-MAR-1995			
: ATTORNEY/AGENT INFORMATION:			
: NAME: STEFFE, ERIC K.			
: REGISTRATION NUMBER: 36,668			

```

? REFERENCE/DOCKET NUMBER: 1488.1140002/EKS/KLM
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 202-371-2600
? TELEFAX: 202-371-2540
? INFORMATION FOR SEQ ID NO: 1:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 1713 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: both
? TOPOLOGY: both
? MOLECULE TYPE: CDNA
? FEATURE:
? NAME/KEY: CDS
? LOCATION: 116..1003
? US-08-467-947A-1

```

Query Match	21.1%;	Score 199.2;	DB 3;	Length 1713;
Best Local Similarity	52.3%;	Pred. No. 2.1e-52;		
Matches 441;	Conservative 0;	Mismatches 403;	Indels 0;	Gaps 0;

QY	13	aaatgacagcatctacaaaggattatctgcgttggtcttcagatgcagcttgcgtgga / 2
Db	136	AAATGACGACATGGGTCACAGAGTTCTCTCTACTGTGGATTTTCTCTGGGCCAAGGATTCA 195
OY	73	gtttccacctctgtgtgtctctcttgattctcttaacctgacatctttgccaatctgac 132
Db	196	GATGTCCTCTTTTGGGCTTCTCTCCGTGTTCTATGTCCTTACACCTCGTGGGGAATGGAC 255
OY	133	catattcttagtgcacggcctggaacacaactcatcaccccaatgattttttcttac 192
Db	256	CATCTTGGGGCTCATCTTCACGTGACCTCCAGCTCCACACCCCAATGTACTTCTCTCTC 315
OY	193	caattatcacctctctggaatctttgaaacacacatgacatgtcccaanaatgtcagtaaa 252
Db	316	ACACCTGGCCGTCGTCAACATCGGCTATGCTCGTGAACACAGATGCCCAATGGCGTGAA 375
OY	253	tttatgacgacatcagaaagaatcagatgatlcatgctgtgcgtgtgtagccagctttcatatt 312
Db	376	CCTCTCGCATCCAGCCAGACCCCATCTCTTTGGTGGTTGATGACATAGACTTTTCTCTT 435
OY	313	ctctgacctgtggagcctactgataatctctctccgcgtgaatgaccttgattgtgtgc 372
Db	436	TTTGAGTTTGCACATACATATGATATCCCTCTCTGTTGTGCTGATGTCACATAGGTAAGT 495
OY	373	agctattgttcggccctcccatctactcaagtatcatgaacagaagatcttgcctccagtc 432
Db	496	GGCCATTCGCCACACCTCTCCGATATTTTCATATATATGACCTGGAAAGTCTGCATCACTCT 555
OY	433	ggcagctgcacccctggtgttaactgtttttagtaactcagttgtgtgtctacacctgcact 492
Db	556	GGGACATCACTTCGTGGAAATGTGGCTCCCTCTGGATATGTGTCATGTGAAGCCATCACTCT 615
OY	493	ccagctgcacctctgtgaaccccatatgtatagatcatctctctgtgaagtcacctgact 552
Db	616	AAGACTGGCCCTTTGTGGGCTCGCTGAAATCAACCACTTCTTGTGTAATATCCGTCTGT 675
OY	553	gctcaagtatctgtgtgttgagaacaaacagcaaatgaaagcgtgaactatctcttgcagtga 612
Db	676	CCTCAGGCTGGCTGTGTGTATACCTGGCTCAACCAAGTGGGTCACTTTTGAAGCCGTGCAT 735
OY	613	gctctccatctaataccctcgacacatcatctccatatcatatgcttttatgttcagac 672
Db	736	GTTCACTCTGTGGTGGACCACTCTTGCTGTGCTGTCTCTACTACACATCTCTGGGGGG 795
OY	673	agatattgagaataagctctgtgaaggtcgacaanaaagcatattggaacatgtgtgttccca 732
Db	796	CATCTGTAGATCCAGTCTGGGGAGGGCGGAGAAAGGCCCTTCCACACGTCTCTCCCA 855
OY	733	ctcaattgtgtgtctcttttttatagtaacagccgctctcgtgtactctgaaccaacttc 792
Db	856	CCTCGCCTATGTGGACCTCTTTTGSAGGGCCATGTCATGATACATGGCCCCCTAAGTC 915

Oy	793	gcccgcctccacggagcagaagaatcggttcttcctcatctgaataatcgaccocat	852
Dd	916	cgccgatccttgtaggacgacagaaagctcttttcttatatttaaacgtttccaacccc	975
Oy	853	gctcg 856 Db 976 gatg 979	
RESULT 7			
US-08-748-506-5			
; Sequence 5, Application US/08748506			
; Patent No. 6159707			
GENERAL INFORMATION:			
APPLICANT: Konnert et al.			
TITLE OF INVENTION: NOVEL SPERM RECEPTORS			
NUMBER OF SEQUENCES: 31			
CORRESPONDENCE ADDRESS:			
ADDRESSEE: Leydig, Volt & Mayer, Ltd.			
STREET: Two Prudential Plaza, Suite 4900			
CITY: Chicago			
STATE: IL			
COUNTRY: US			
ZIP: 60601-6780			
COMPUTER READABLE FORM:			
MEDIUM TYPE: Floppy disk			
COMPUTER: IBM PC compatible			
OPERATING SYSTEM: PC-DOS/MS-DOS			
SOFTWARE: Patentin Release #1.0, Version #1.25			
CURRENT APPLICATION DATA:			
APPLICATION NUMBER: US/08/748,506			
FILING DATE: 08-NOV-1996			
CLASSIFICATION: 435			
PRIOR APPLICATION DATA:			
APPLICATION NUMBER: US 60/033,751			
FILING DATE: 09-NOV-1995			
CLASSIFICATION: 435			
ATTORNEY/AGENT INFORMATION:			
REFERENCE/DOCKET NUMBER: 74940			
TELECOMMUNICATION INFORMATION:			
TELEPHONE: 312-616-5600			
TELEFAX: 312-616-5700			
INFORMATION FOR SEQ ID NO.: 5:			
SEQUENCE CHARACTERISTICS:			
LENGTH: 966 base pairs			
TYPE: nucleic acid			
STRANDEDNESS: double			
TOPOLOGY: linear			
MOLECULE TYPE: DNA (genomic)			
US-08-748-506-5			
Query Match 21.1%; Score 198.8; DB 3; Length 966;			
Best Local Similarity 51.9%; Pred. No. 2.1e-52;			
Matches 446; Conservative 0; Mismatches 411; Indels 0; Gaps 0.			
OY	50	tcttcagactcgacctggtcgtggagtcttcacatctcctcgttgttcttcaattcttacact	109
Dd	67	ttctctTGAGGTCCTGGAGAATGCCTCCCGTTCAACCTCATCTTGTACTGTTCTTA	126
OY	110	gtgaccacatttggtgacaatcgcacatlattctagtgtcaaggcctggaaccaacttcac	169
Dd	127	GATACATAACAGAAATAAATCTCATAGTCCCTTGCTATTGTGACCAGNCACATCTCAC	186
OY	170	accaccatgatatctttcttaccacatctatcacctcctcgtgaactcttgttaaccacacagt	229
Dd	187	ACCCCCATGTACTTCTTCTGGCCAACCTGTCTCTCCGGAATGGCTATATCTTGCTCT	246
OY	230	acagtcacacaatgctatgtaaattatcgacgacatcagaaagatcatcaattatcgttgc	289
Dd	247	GTCATATCCCAGATNCTGCAGAGCCTTGAGAGTAGGCCAGAGAAATCTCTGGAGAGGA	306
OY	290	tgttgagccagctttcataattcttgaccttgagggtgactgaatatcttccctggcc	349

Db	307	TTGTGGCACACGAGTCTTTTGTGGCATTTTTGGTATTAAGTGGTCTGCTTATTGGCA	366
Qy	350	gtlcatgcttccttgatgtgttctgtagctatctgtgcgcctcccaactcaagtatcatg	409
Db	367	GCCATGGGCTTTGACCGCTGGATGGCCATATGCTCCCACTGCACATVGCMAACCCGAATG	428
Qy	410	caaccagagacctgcctcccaatgtggcagctgcataccctgggttaetgtttagtaetca	465
Db	427	AGTCNMGAGGATGATGGCCATTGGCAATTTGTTTCATGGGGAAAGGATGCATAGTAAGT	486
Qy	470	gtgtgtgtgtctaacctgacctcccaagctgcacctctgtgacccctatgtgatatcatc	529
Db	487	CTGGGACAAACCAATTTTATTTTTCTCCTTGAACCTTCTGTGACCCCTGTGAATATGACAC	546
Qy	530	ttctctctgtgaagtcctcctgcactgtcctaagttatcttgtgtgtgagacaagcaatgag	589
Db	547	TTCTCTGTGACCTTGACACCTCTCTGCACTTGCCTGTGAGATATCATCCCAAAAGAG	606
Qy	590	gctgaacattcctctgtcatgtgagctcttcacatcataaccctgcgaactcattcctata	649
Db	607	GCTGCACTTTTGTGTGTGACAGTCTCTGCAATATCTAGCCCATTTTGTGATCATTTAT	666
Qy	650	tcaatgtctttatgtgtccgaagagiatgtgagatacagtcgtgtgaagtcgacaataa	709
Db	667	TTCTATGTCAAAATTTCTATTGTGCAAGTGTCTGATGCTTACCTTGAGGGGCCCATATAA	726
Qy	710	gcattgtgagacatgtgtgtcccatctaattgtgtgtctcttllttaagtaacagccgtc	769
Db	727	GCTTTTCCACCTGTGTGCTGCACCTACTTGTAGTCAACATTTTATTTGCGTCAGCATGT	786
Qy	770	tctgtgtactcgtcaaccaacctgtgcacctccaagtaaccaagaagaatggtttctctc	829
Db	787	ATTACCTTTTGTGAGGCCCAAGTCTAGCCACTCACACAGGAATGACAAATTTCTGGCCCTC	846
Qy	830	ttctatgtaatcatgtgaccccatgtaatcccttatatacaacttagaacaagaag	889
Db	847	TTCTACACAGTAGTACATTCATCTGTGACCTGTATCATCTTATAGTTTAAAGAACAGGA	906
Qy	890	gtaagaagaagctttaaag 909	
Db	907	GTCAAGGACGACGATGAGAG 926	
RESULT 8			
US-08-748-506-6			
: Sequence 6, Application US/08748506			
: Patent No. 6159707			
: GENERAL INFORMATION:			
: APPLICANT: Ronnelt et al.			
: TITLE OF INVENTION: NOVEL SPERM RECEPTORS			
: NUMBER OF SEQUENCES: 31			
: CORRESPONDENCE ADDRESS:			
: ADDRESSEE: Leydig, Volt & Mayer, Ltd.			
: STREET: Two Prudential Plaza, Suite 4900			
: CITY: Chicago			
: STATE: IL			
: COUNTRY: US			
: ZIP: 60601-6780			
: COMPUTER READABLE FORM:			
: MEDIUM TYPE: Floppy disk			
: COMPUTER: IBM PC compatible			
: OPERATING SYSTEM: PC-DOS/MS-DOS			
: SOFTWARE: PatentIn Release #1.0, Version #1.25			
: CURRENT APPLICATION DATA:			
: APPLICATION NUMBER: US/08/748,506			
: FILING DATE: 08-NOV-1996			
: CLASSIFICATION: 435			
: PRIOR APPLICATION DATA:			
: APPLICATION NUMBER: US 60/033,751			
: FILING DATE: 09-NOV-1995			
: CLASSIFICATION: 435			
: ATTORNEY/AGENT INFORMATION:			

```

; REFERENCE/DOCKET NUMBER: 74940
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-616-5600
; TELEFAX: 312-616-5700
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 966 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-748-506-6

```

```

Query Match          20.4%; Score 192; DB 3; Length 966;
Best Local Similarity 53.5%; Pred. No. 2.8e-50;
Matches 399; Conservative 0; Mismatches 347; Indels 0; Gaps 0;

```

```

QY 164 ctccataccccatgtatcttttcttcaaatcctcctgactcttggtaacc 223
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 181 CTACACACCCCGCATGCTTCTTCTGCGCAACTGTCTCCTCGAGATGGCTATAC 240
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 224 acatgacagtcacacaaagctagtaattatgacagatcaggaagaatagttat 283
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 241 TGCTCTGTCAATCCCAAGATGCTGCAAGCCTTGAGTGAGGCCAGAGATCTTAG 300
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 284 cgtgctgtgtagccagctttcatacttctgacctgggggactagataatcttc 343
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 301 GAGGAGATGTCACACAGATGTTTCTTCATATTTTGGCATAACTAGTGTGCTTA 360
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 344 ctggcgctatgctcttgattgtttgtgactatcttgccgctccctactacgtt 403
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 361 TTGGCACCACATGGCTTTGACCGCTGCATGGGCAATGCTCCCACTCACTAGCA 420
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 404 atcagcaccagagacactcctccagctgagcagctgacccgtgactagtttagt 463
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 421 CGATAGATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 480
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 464 aactcagtggtgtgtctaacctgactcctcagctgacactgtagacccctatg 523
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 481 GTAGGTCGTGGACAGACCAATTTNATTTNCTTCTTGAACCTTCTGAGACCT 540
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 524 gatcactctctgtgtaagccctcgcactgctcaagtatcttggctgtagaaca 583
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 541 GACGACTTCTTCTGTGACCTTCAACCTTCTGCTGCTGCTGCTGCTGCTGCT 600
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 584 aatgagctgaactatcctctgtcagtgagctcttccatcaatacccccagcact 643
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 601 AAGCAGGCTGCACTTGTGGGAGCAATCTCTGTATATCTATCTATCTATCT 660
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 644 ctatatacatgctcttattatgtcgaagcagatgagatagacagctgctgaag 703
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 661 CTATATCTTATGTCAAGATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 704 caaaaagcattggagacatgtgttccactaatltgtggtctctttttatagta 763
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 721 CATTAACCTCTTTTCCACTGTCTCTCAOACTACTTGTAGTCACACACTCTTA 780
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 764 ggcgtctctgtgtaaccgcaacacctgcgccagctccaaagcaagaagaatg 823
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 781 GTGTCCTTTACCTTTGAGGCCCCAAGTCTAGCCACTCACAGGAATGGAACAC 840
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 824 tctctcttctatggaatcattgcaaccatgctgaatcccttatatacttaga 883
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 841 GCCCTCTTTACACAGCAGTACATCTGTAACCTTATCATCTACAGCTTAAG 900
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 884 aagagagtaagaagaagctttaaag 909
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 901 AAGGAAGTCAAGGACGACCTGAGAG 926
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 9
US-08-748-506-9

```

; Sequence 9, Application US/08748506
; Patent No. 6159707
; GENERAL INFORMATION:
; APPLICANT: Ronnett et al.
; TITLE OF INVENTION: NOVEL SPERM RECEPTORS
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Leydig, Volt & Mayer, Ltd.
; STREET: Two Prudential Plaza, Suite 4900
; CITY: Chicago
; STATE: IL
; COUNTRY: US
; ZIP: 60601-6780
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/748,506
; FILING DATE: 08-NOV-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/033,751
; FILING DATE: 09-NOV-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; REFERENCE/DOCKET NUMBER: 74940
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-616-5700
; TELEFAX: 312-616-5600
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 984 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-748-506-9

```

```

Query Match          19.6%; Score 185.2; DB 3; Length 984;
Best Local Similarity 50.7%; Pred. No. 3.7e-48;
Matches 445; Conservative 0; Mismatches 433; Indels 0; Gaps 0;

```

```

QY 14 aatagacatcatcacaagagattatctcgtggttctcagaatcagactgagctgag 73
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 31 AATGGACTTGTGCTCGAATTCATCTTGAGGGATACCTGTGGCCGACACTGTAAG 90
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 74 ttccactccttggtctcttctgattcttaacactgtgacacacttggcaactgac 133
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 91 ATTCCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 150
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 134 attattagtgtaacgctggagacacaaattatcccaatgattttttcttacc 193
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 151 ATTAATTAACATCACTGTGAGGACGAGCTGAGAGGCCAATGATCTTTTCTGAG 210
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 194 aatcctacactcctgtaacttctgttaacacacatgtaacagtcacaaatgtag 253
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 211 ACCTTCTCTTTTGTGGAGTGTGTTTATTAACATGCTGTAATCCCAAGCTTCA 270
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 254 ttatgacagcatcagaagaatatacagttatcggtgtgtgtagccagctttcata 313
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 271 ATTCGTGACAGGAGGCAAAAGATTCCTTTGGGGTGTCTTCTACAGGCCCTTG 330
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 314 ctggccttgagggtcagtaaatcctctcctgagcgctgaatgcttgaattggtgta 373
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 331 CTGTGCTGGGGGCAACAGGTTTTTCTTTTGTGCTGCTGCTGCTGCTGCTGCT 390
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 374 gctattgctgagcctctccatctacatcagttatcagtcacagaagactgctccag 433
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 391 GCCATCTGCAAAACCTTACATTAATCAACCATATGAGGCCAAGAGATGTCTTCT 450
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

QY 434 gcagctcacctcgtggttactgtttagtaactcagctgtgtgtctacccctgactctc 493
DB 451 GTTACTCTCTGTTTATTTTGGGCTTCTCTTCATGGCCAGTCCAGTTGATGCTTTCC 510
QY 494 cagctgcacactctgtgacccctatgtagatcactctctctgtgagtcctccgactg 553
DB 511 AAGACATTTCATCTGTGGTCCAAACATTAATTCCTACCTTTTCTGTGATTTTGACACTG 570
QY 554 ctcaagttatctgtgttgagacacagcaaatgagctgaactactctctgtcagtgag 613
DB 571 GCAAAATCTTCCTGTCGAGAAACAGAGCTATGAGATGCTGTTTATACCTTGCTGTA 630
QY 614 ctcttcacatcaataaccctcgacactcactctatcactatgcttattgtccgagca 673
DB 631 ATGTGCTTTTGGCTTCTTTTATAGCCATCTTGTGATACAGCAATATATGATGACCC 690
QY 674 gttatgagatatacagctctgagagtcgacaaagacattggagatgtgttccat 733
DB 691 ATATGTGACTCCCTTATAGCCAGGAGGAGCAGACAGAGCTTTTTCACCTGCTCTCAT 750
QY 734 ctatgtgtgtctctcttattatagtagacagcgctctctgtgtacactgcaaccacttg 793
DB 751 CTCATGTCTCCCTCTCTAATGATGAGCAGCTGTGATTTATATACGAGCCAAAGCAG 810
QY 794 cccagctcccaagagcaagaagaatgttctctctctctcattgtaacatcagcccatg 853
DB 811 AGAAGCAGAGTGGACCAACAGAGAGGCTGCTTGTGAACATGATGTGACACCCCTT 870
QY 854 ctgagacccttatatacacttagaagaagaagt 891
DB 871 CTGACCCCTGTATCTACACCTGCGCAACAGAGGT 908

RESULT 10

US-09-439-313-526
; Sequence 526, Application US/09439313
; Patent No. 6329505
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, David C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan Louise
; APPLICANT: Jiang Yuxi
; APPLICANT: Reed, Steven G.
; APPLICANT: Kalos, Michael
; APPLICANT: Fanger, Gary
; APPLICANT: Retter, Mark
; APPLICANT: Solk, John
; APPLICANT: Day, Craig
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; FILE REFERENCE: 210121.427C9
; CURRENT APPLICATION NUMBER: US/09/439,313
; CURRENT FILING DATE: 1999-11-12
; NUMBER OF SEQ ID NOS: 575
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 526
; LENGTH: 963
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-439-313-526

Query Match

Best Local Similarity 49.1%; Score 120; DB 4; Length 963;
Pred. No. 7.9e-28;
Matches 318; Conservative 0; Mismatches 330; Indels 0; Gaps 0;

QY 110 gtaacatcttggcaactgacatattctagtgacagcctcggaaccaatctcat 169
DB 106 gttgcaatgtttggaactgctgtctctacatgtaaggaagcaagcagctgac 165
QY 170 accccatgatttttcttacaatcattactcctgtgactctgttttacaacatgtc 229

DB 166 gctccgatgtacctcttctctgcatctgtgcagcaattgaacctgacctatccacatcc 225
QY 230 acagttcccaacaatgtcagtaaatlttagcagcactcaggaagaatgaattacgtgtgc 289
DB 226 accatgtcctaagatcccttgcctcttctgtttagttatcccgagagatgaactttagggcc 285
QY 290 tgtgtagccaagcttctcaattctctgtctgtggggtactgaatatctctctggcc 349
DB 286 tgtcttccagagatgtcttattatcagtcctcctcagcaattgaatccacatctgtg 345
QY 350 gttatgctttgattgtgttagatattgtgtgctctcctcctcaattacatg 409
DB 346 gcaatgaccttgacgcttatgtgtgcatctgcaaccactgtgcactgtgtgcgtc 405
QY 410 caccagagactctgctccagttgtgcagctgcacacctgtgttactgttttagtaactca 469
DB 406 aacatacagtaaacagccagattgtgacatcggtgtgtgtgtccgagatccctctt 465
QY 470 gttgtgtgtctacacctgactctcagctcgaactcgtgacatgacctatgtatagatcac 529
DB 466 ttcccaactgctctctgtgatacagcgctgtgctctcgcactcctcaatgtcctcgcac 525
QY 530 ttctctgtgaagtcctctgacatgctcaagttaactgtgtgtgagacaagcaaatgag 589
DB 526 tccatgtgtgtccacagagatgtaagtgtgctctatgcagacacttggccaagtgt 585
QY 590 gctgaactatctctgtcagtgagctcttccatcattacacccctgaacatcactatata 649
DB 586 gtaatgtcttactgcatcttctgtctatgtggcggtgagcgaatgttcatctctgt 645
QY 650 tcatatgcttattatgtcagagcagatgtgaagatagactgtctgaaggtgcagaaaa 709
DB 646 tccatttctgataaacagacagctgtctgaactgtccttccaaagtcagagcgcggaag 705

RESULT 11

US-08-465-980-1
; Sequence 1, Application US/08465980
; Patent No. 5756309
; GENERAL INFORMATION:
; APPLICANT: Soppet, Daniel R.
; APPLICANT: Li, Yi
; APPLICANT: Rosen, Craig A.
; APPLICANT: Ruben, Steven M.
; TITLE OF INVENTION: HUMAN G-PROTEIN RECEPTOR HPR4J70
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CARELIA, BYRNE, BAIN, GILFILLAN, CECCHI,
; ADDRESSEE: STUART & OLSTEIN
; STREET: 6 Becker Farm Road
; City: Roseland
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07068
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/465,980
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Ferraro, Gregory D.
; REGISTRATION NUMBER: 36,134
; REFERENCE/DOCKET NUMBER: 325800-446
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-994-1700


```
TELEFAX: 201-994-1744
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1474 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 274..1233
US-08-465-980-1

Query Match      12.4%; Score 116.8; DB 1; Length 1474;
Best Local Similarity 48.8%; Pred. No. 9.6e-27;
Matches 316; Conservative 0; Mismatches 332; Indels 0; Gaps 0;

QY 110 gtgacacatcttggaacatcgacattatctagtgtcaagcctgagacacaaattcat 169
Db 379 GTGGCAATGTGTGGAACCTGCATCGTGTCTCATCGTAAGAGACGGAAACGACGCTGCAC 438
QY 170 accccatgatttttcttaccacatctacacccctgagatcttgtaacacacatgt 229
Db 439 GCTCCGATGACTCTTTCTCTGATGCTTGACAGCCATTACCTGGCCTTATCCACATCC 498
QY 230 acagtcaccaaatgtagtaaatatagcagcatcaagaagaatlaacagttatcgtgac 289
Db 499 ACCATGCTTAAGATCGTTGGCCCTTTCTGTGATTGCCAGAGATTACATTTGAGGCC 558
QY 290 tctgttagccagcttttcaattcttgagccttgagggtctctgaatacttcctcgcc 349
Db 559 TGCTTTACCCAGATGTTCTTTATTCATGCCCCCTTCAGCCATTGAATCCACCATCTCTCG 618
QY 350 gtaatgctcttgatgtgttgatgactatgtgcgctctcactacactacatgatactg 409
Db 619 GCCATGGCCCTTGAACGTTATGTGGCCATCTGCACACCCACTGGGCCATGTGACGTCTC 678
QY 410 caccagagactcgtcctcagttgagcagctgcatcctggttactggttttagtaactca 469
Db 679 AACAAATACATTAACAGCCGATTTGGCATGTGTGTGGCTGGGATCCCTCTTTT 738
QY 470 gttgtgttgctacacccctgacactccagctgacactgctgagacccctatgtagatcac 529
Db 739 TTCCCACTGCTCTGCTGTGATCAAGCGGCTGGCCTTTGCCCACATCAATGTCCTCGCAC 798
QY 530 ttctctctgtaagtcctgacactgctcaagtatactgtgttgagacaacagcaaatgag 589
Db 799 TCTATTGTGTCCACGAGATGTAATGAATGTGGCCATATGACACATTTGCCCAATGTG 858
QY 590 gctgaactatctctgtcagtgagctctccactcaatataccctgacatcactcattata 649
Db 859 GATATAGGTCTTACTGCGCATTTGCTGTGTCATGGGCGCTGACGTAATGTTGATCTCTT 918
QY 650 tcatatgctttatgtgcgagcagatagagatacagatcgtctgaagtgtagaataaaa 709
Db 919 TCTATTATTTGTGATATACAGAGGTTTCTGCAACTGCCCTTCAAGTAGAGCGGGCCACG 978
QY 710 gcaattggagacatgtgttccactcaatgtatgtgtctctttttat 757
Db 979 GCCTTTGGAACTGTGTGTGCACACATTTGTTGTGTACTGCCCTTCTAT 1026

RESULT 12
US-09-053-303-1
; Sequence No. 1, Application US/09053303
; Patent No. 5948890
; GENERAL INFORMATION:
; APPLICANT: Soppe, Daniel R.
; APPLICANT: Li, Yi
; APPLICANT: Rosen, Craig A.
; APPLICANT: Ruben, Steven M.
; TITLE OF INVENTION: HUMAN G-PROTEIN RECEPTOR HPRAL70
```

```
NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN, CECCHI,
; ADDRESSEE: STUART & OLSTEIN
; STREET: 6 Becker Farm Road
; CITY: Roseland
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07068
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/053,303
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/465,980
; FILING DATE: 06-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Ferraro, Gregory D.
; REGISTRATION NUMBER: 36,134
; REFERENCE/DOCKET NUMBER: 325800-446
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-994-1700
; TELEFAX: 201-994-1744
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1474 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 274..1233
US-09-053-303-1

Query Match      12.4%; Score 116.8; DB 2; Length 1474;
Best Local Similarity 48.8%; Pred. No. 9.6e-27;
Matches 316; Conservative 0; Mismatches 332; Indels 0; Gaps 0;

QY 110 gtgacacatcttggaacatctgacacattatctagtgtcaagcctgagacacaaattcat 169
Db 379 GTGGCAATGTGTGGAACCTGCATCGTGTCTCATCGTAAGAGAGAACGACGCTGCAC 438
QY 170 accccatgatttttcttaccacatctacacccctgagatcttgtaacacacatgt 229
Db 439 GCTCCGATGACTCTTTCTCTGATGCTTGACAGCCATTACCTGGCCTTATCCACATCC 498
QY 230 acagtcaccaaatgtagtaaatatagcagcatcaagaagaatlaacagttatcgtgac 289
Db 499 ACCATGCTTAAGATCGTTGGCCCTTTCTGTGATTGCCAGAGATTACATTTGAGGCC 558
QY 290 tctgttagccagcttttcaattcttgagccttgagggtctctgaatacttcctcgcc 349
Db 559 TGCTTTACCCAGATGTTCTTTATTCATGCCCCCTTCAGCCATTGAATCCACCATCTCTG 618
QY 350 gtaatgctcttgatgtgttgatgactatgtgcgctctcactacactacatgatactg 409
Db 619 GCCATGGCCCTTGAACGTTATGTGGCCATCTGCACACCCACTGGGCCATGTGACGTCTC 678
QY 410 caccagagactcgtcctcagttgagcagctgcatcctggttactggttttagtaactca 469
Db 679 AACAAATACATTAACAGCCGATTTGGCATGTGTGTGGCTGGGATCCCTCTTTT 738
QY 470 gttgtgttgctacacccctgacactccagctgacactgctgagacccctatgtagatcac 529
Db 739 TTCCCACTGCTCTGCTGTGATCAAGCGGCTGGCCTTTGCCCACATCAATGTCCTCGCAC 798
```


FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0441 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-855-0555
TELEFAX: 650-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1828 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: PROSNOT01
CLONE: 364702
US-08-988-876-2

Query Match 10.2%; Score 96; DB 3; Length 1828;
Best Local Similarity 55.7%; Pred. No. 3.2e-20;
Matches 206; Conservative 0; Mismatches 160; Indels 4; Gaps 1;
QY 35 ttatttctgtggttctcgaatcgaacattgctgagttccactcctgtgtcttc 94
DB 814 TTCCTCTCTAGAGCTCAGGGGATCAGAACTGCAGCCAGCTGTGCTGGCTGTTC 873
QY 95 ttgattcttaactgacatcttggcaatcgaacattatctagtgtcaagcctg 154
DB 874 CTGCTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 933
QY 155 gacaccaactcaatacccatgtattttcttcaacatcaactcactcctgattc 214
DB 934 GACTCCACCTCCACACCCCATGTACTTCTCTCTCAACCTGCTGCTGCTGCTGCT 993
QY 215 ttttaaccacatgttaagtcacaaatagtaattatcagcactaagaaagta 274
DB 994 GGTTCACCTCCACACGAGTCCCAAGATGATGTGACATCCAGTCTCAGCAGAGTC 1053
QY 275 atcagttatgctgctgtgagccagctttcatatttctgcttgggggagctacta 333
DB 1054 ATTCCTCATGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1113
QY 334 --atacttctcctggcgcgtcatgtccttgaattgttgaactatttgcgcctc 390
DB 1114 GAGAGACATGCTCTGATGTGATGGCTATGACTGGTTGTAGCAATCTGTACCCGCT 1173
QY 391 ccattactca 400
DB 1174 ATATCAATCA 1183

RESULT 15
US-08-599-252-84/c
Sequence 84, Application US/08599252
Patent No. 570343
GENERAL INFORMATION:
APPLICANT: DRAVNA, DENNIS T.
APPLICANT: FEDER, JOHN N.
APPLICANT: GIMKE, ANDREAS
APPLICANT: KIMMEL, BRUCE E.
APPLICANT: THOMAS, WINSTON J.
APPLICANT: WOLFF, ROGER K.
TITLE OF INVENTION: METHOD TO DIAGNOSE HEREDITARY
TITLE OF INVENTION: HEMOCHROMATOSIS
NUMBER OF SEQUENCES: 124
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORRISON & FOERSTER
STREET: 2000 Pennsylvania Ave. N.W., Suite 5500
CITY: Washington
STATE: DC

COUNTRY: USA
ZIP: 20006-1888
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/599,252
FILING DATE: 09-FEB-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: MURASHIGE, KATE H.
REGISTRATION NUMBER: 29,959
REFERENCE/DOCKET NUMBER: 9053-0001.21
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 887-1500
TELEFAX: (202) 887-0763
TELEX: 90-4030
INFORMATION FOR SEQ ID NO: 84:
SEQUENCE CHARACTERISTICS:
LENGTH: 1320 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-599-252-84

Query Match 9.6%; Score 90.2; DB 1; Length 1320;
Best Local Similarity 65.8%; Pred. No. 1.8e-18;
Matches 131; Conservative 0; Mismatches 68; Indels 0; Gaps 0;
QY 719 acatgtgttcccaatgaatgtgtgtctcttttataagtaagccgtctcgtgtac 778
DB 1319 ACTTGACGCTGCATCTATGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1260
QY 779 ctgcaaccacttgcagctccaaggaagaaagatggttctctctctataga 838
DB 1259 CTCGAGCCAGCAAAATCTTATCTCCAGGACCAAGGCAAGTTCTTCTTCTACCA 1200
QY 839 atcattgacccatgctgaatcccttatatacacttaagaaaggaagaaagaa 898
DB 1199 ATTGTCACTCCAGTGTAAACCCCTGATCTATACATCAAGAAAGATGTTAAAG 1140
QY 899 ggccttaaaaggtcgtgtg 917
DB 1139 GCCATGAAGAAGGTGCTG 1121

Search completed: August 25, 2002, 18:57:33
Job time: 4586 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 25, 2002, 13:20:02 ; Search time 1886.27 Seconds

(without alignments)
6747.512 Million cell updates/sec

Title: US-09-800-321A-3

Perfect score: 943
Sequence: 1 tatgaattgggtaaatgaca.....tctcttaatacaagaataaa 943Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

EST: *
1: em_estha: *
2: em_esthm: *
3: em_estln: *
4: em_estmu: *
5: em_estov: *
6: em_estpl: *
7: em_estro: *
8: em_estl: *
9: gb_estl: *
10: gb_est2: *
11: gb_hic: *
12: gb_gss: *
13: em_gss_hum: *
14: em_gss_inv: *
15: em_gss_pln: *
16: em_gss_vtl: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	337.4	35.8	479	12	AQ077154
2	226.6	24.0	2021	11	BC016940
3	204.6	21.7	523	9	A1604386
4	200.2	21.2	692	12	A2090606
5	197.2	20.5	642	12	A2969227
6	193.4	20.5	1394	11	AK017036
7	193.4	20.5	3063	11	AK016560
8	192.6	20.4	632	12	A2511623
9	185.4	19.7	428	10	N68399
10	184	19.5	326	12	A2694245
11	178.6	18.9	640	12	A2516219
12	177.2	18.8	788	12	BH113304
13	176	18.7	853	12	BH331857
14	175.6	18.6	588	12	A2431817
15	173.6	18.4	740	12	A2607393
16	165.8	17.6	679	12	AG167722
17	164.2	17.4	632	12	A2765752

C	18	164	17.4	548	12	A271125	A271125	RPCI-23-1
C	19	163.2	17.3	674	12	AG131684	AG131684	Pan trogl
C	20	159.6	16.9	680	12	A2709687	A2709687	RPCI-24-8
C	21	156.2	16.6	675	12	A2396801	A2396801	IM0161E05
C	22	154.8	16.4	1501	11	AK016338	AK016338	Mus muscu
C	23	154.2	16.4	522	12	A2720443	A2720443	RPCI-24-1
C	24	154.2	16.4	660	12	A2380178	A2380178	IM0136A04
C	25	154	16.3	686	12	A2086625	A2086625	RPCI-23-2
C	26	153.8	16.3	534	12	A2312830	A2312830	IM0028J19
C	27	153.8	16.3	646	12	A2638594	A2638594	IM0498M06
C	28	151.8	16.1	608	12	A2112932	A2112932	RPCI-23-1
C	29	151	16.0	628	12	A2977433	A2977433	2M0233P17
C	30	150.2	15.9	797	12	A2909618	A2909618	RPCI-24-2
C	31	150	15.9	508	12	A2373636	A2373636	IM0126C20
C	32	149.6	15.9	484	12	AQ310400	AQ310400	CITBI-E1-
C	33	149	15.8	639	12	A2418543	A2418543	IM0194J10
C	34	148.2	15.7	649	12	BH067870	BH067870	RPCI-24-3
C	35	147.6	15.7	605	12	A2642411	A2642411	IM0505I22
C	36	147.4	15.6	532	12	A2519123	A2519123	RPCI-11-6
C	37	147.2	15.6	553	12	AQ482125	AQ482125	RPCI-11-2
C	38	147.2	15.6	762	10	BG193339	BG193339	RST12467
C	39	147	15.6	680	12	A2235331	A2235331	RPCI-23-6
C	40	146.8	15.6	534	12	A2664867	A2664867	IM0545E05
C	41	146.4	15.5	632	12	A2382351	A2382351	IM0154L21
C	42	146.2	15.5	632	12	A2382351	A2382351	IM0154L21
C	43	146.2	15.5	796	10	BG197640	BG197640	RST17016
C	44	146	15.5	619	9	A1148854	A1148854	qc70a02.x
C	45	145.8	15.5	678	12	A2964406	A2964406	2M0234003

ALIGNMENTS

RESULT 1
LOCUS AQ077154/c
DEFINITION CIT-HSP-2354D1.TF CIT-HSP Homo sapiens genomic clone 2354D1, DNA
sequence.
ACCESSION AQ077154
VERSION AQ077154.1 GI:3438338
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 479)
AUTHORS Adams,M.D., Rounsley,S.D., Zhao,S., Baas,S., Linher,K., Golden,K.,
Berry,K., Granger,D., Sub,E., Wible,C., Shizuya,H., Simon,M. and
Venter,J.C.
TITLE Use of a random human BAC End Sequence Database for Sequence-Ready
Map Building
JOURNAL Unpublished (1998)
COMMENT Other GSSs: CIT-HSP-2354D1.TR
Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdadams@tigr.org
Clones are available from Research Genetics (info@resgen.com). BAC
end search page:
http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
Seq primer: M13-21
Class: BAC ends.

FEATURES

SOURCE

location/Qualifiers
1..479
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="2354D1"
/clone_lib="CIT-HSP"
/sex="Male"
/cell_type="Sperm"

/note="Vector: pBelOBAC11; Site_1: HindIII; Site_2: HindIII"

BASE COUNT 139 a 100 c 111 g 129 t

ORIGIN

Query Match 35.8%; Score 337.4; DB 12; Length 479;
Best Local Similarity 90.9%; Pred. No. 5.8e-81;
Matches 370; Conservative 0; Mismatches 36; Indels 1; Gaps 1;

Qy 538 tgaagccctcctcagctcagctatctctgtgtgagacaacagcaatgagctgaact 597
Dy 479 TGAAGCCCTCTCTCTGCTCAAGTGTCTGTGTGACACACATGCAATGAGCTGACT 420

Qy 598 attccttgtagtgagctctcctcaataaccctgagactcctcctatataatgac 657
Dy 419 ATTCTTCATGAGTGTGCTATTCCTTAATACCGGTGACACTCAACCTTATGCTATGC 360

Qy 658 tttaattgctcagagatattgagagataagctcgtcgtgagagtcgacaaaagcaattgg 717
Dy 359 TTTTATGTTCCACAGCAGTGTGAGATGCCACTGCTGACAGCAACGAAGCATTTGG 300

Qy 718 gacatgtgttcctcctcaatctggtgtgtctctttttatagtagcagcgtctgtgta 777
Dy 299 GACATGTGCTCCATCTCAATGTGTGTGTCACCTTTTATGTACAGTATCTCCATGTA 240

Qy 778 cctgcaacacacttcgcagcagctcacaaggaagaagaatgttctcctcctatg 837
Dy 239 CCTGCAACACCTTCACCTCAGCTCCAAAGACCGGGGAAGATGGTTCTCTCTGTGG 180

Qy 838 aatcattgcaacactgtcgtatcccttatatactatgagacaagaaggttaagga 897
Dy 179 AATCATTTGACCCATGCTGAATCCCTTATATATACACTTGTGGAACAAGAGGTAAAGGA 120

Qy 898 aggccttaaaaggtgtgtgtgc-aagagctcctccttaataagaataaa 943
Dy 119 AGCCTTTAAAGGTGTGTGTGCAAGAGCTCTTATATCAAGAATAAA 73

RESULT 2

BC016940 2021 bp mRNA linear HTC 09-NOV-2001

LOCUS

DEFINITION Homo sapiens, similar to olfactory receptor, family 2, subfamily A, member 4, clone IMAGE:4424116, mRNA.

ACCESSION BC016940

VERSION BC016940.1 GI:16877381

KEYWORDS HTC.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 2021)

AUTHORS Strausberg, R.

TITLE Direct Submission

JOURNAL Submitted (05-NOV-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

CONTACT: MGC help desk

EMAIL: cgabs-remail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305

Web site: <http://www-shgc.stanford.edu>

Contact: (Dickson, Mark) mdg@paxill.stanford.edu

Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.llnl.gov>

Series: IRAC Plate: 27 Row: 1 Column: 22

This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis, Genomescan gene prediction, similarity but not identity to protein

This clone has the following problem: frame shifted.

FEATURES

source

Location/Qualifiers

1. 2021

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:4424116"

/tissue_type="kidney, hypernephroma"

/clone_lib="NIH_MGC_89"

/lab_host="DH10B"

/note="Vector: PCMV-SPORT6"

BASE COUNT 470 a 546 c 427 g 578 t

ORIGIN

Query Match 24.0%; Score 226.6; DB 11; Length 2021;
Best Local Similarity 53.3%; Pred. No. 1.2e-50;
Matches 478; Conservative 0; Mismatches 419; Indels 0; Gaps 0;

Qy 13 aaatgacagcatcacaagagttatctcgtggttccagatgacttgctgga 72
Dy 437 AAATGACAAATGTCACAGATGCTCTACTGAGATTCTCTGGGCCCAAGATTCA 496

Qy 73 gttccacactctgtgctcctcgtatctcctcaactcgtgacacatcgtcagac 132
Dy 497 GATGCTCTCTTTGGGCTCTTCCCTGTCTATGTCTTACACCTGCTGGGGAATGGAGAC 556

Qy 133 cattattctagtgtaacgctcgtgacaaccaactcaactcaatgtatltttcttac 192
Dy 557 CATCTGGGGCTCATCTACTGACTCCAGACTCCACACCCCATGATCTTCTCTCTC 616

Qy 193 caactatcactcctcgtatcttcttgaacacacatgatacctccacaatgctagtaa 252
Dy 617 ACACCTGGCGCTGTCACATGCTGATGCTGCAACACATGCTGCCGATGCTGGTGA 676

Qy 253 ttatgacagcatcagaagaatcagttatcgtgctgctgtgtagccagcttcatat 312
Dy 677 CCTCTGATGATCAGCAAGCCCATCTCTTGTGGGTGACATGATGACCTTCTCTT 736

Qy 313 tctgaccttgaggactactgaatctctcctcgtgcgctcgtcgtcgttctggtgt 372
Dy 737 TTTGAGTTTGCACATATGATGATGCTCTCTGTGTGTGATGCTGATGCTGATGCT 796

Qy 373 agctattgtgctcctcctcctcaactcaactcaactcaactcaactcaactcaact 432
Dy 797 GGCCATCTGCAACCTCTCCCATTTTATCATCATGACCTGGAAAGTCTCATCATCT 856

Qy 433 ggcagctgcatcctcgtggttactggttttagtaactcagtggtgtgtctacact 492
Dy 857 GGCCATCTGCAACCTCTCCCATTTTATCATCATGACCTGGAAAGTCTCATCATCT 916

Qy 493 ccagctgcacactcgtgacccctatgatatgaatcaacttctcgtgtagagctcact 552
Dy 917 AAGACTGCCCTTTGTGTGGGCTCTGGAATCAACACTCTCTGTGAATCCGTCTGT 976

Qy 553 gctcaagttatctcgtgtgtgtgagacaacagcaatgagctgaactcctcgtcagga 612
Dy 977 CCTAGGCTGGCTGT 1036

Qy 613 gctctcactaataaccctcgtgacactcactcactcactcactcactcactcactc 672
Dy 1037 GTTCATCTGTGTGGACACATCTGCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1096

Qy 673 agtattgagatacagctcgtcgtgagagtcgacaaaagcaatttgagagatgttccca 732
Dy 1097 CATCTGAGATCCAGCTGTGGGAGGGCCGAGAAAGGCTTCTCCACATCTCTCTCCCA 1156

Qy 733 tctaattgtgtcctcttttttagtaagcagcgtctcgtgtgacacacactc 792

Dd	1157	COTGCGCTAGTGGACATCTTTTGGAAGGCCCATTCGTCAATGTACATGCAGGCCCTTAAGTC	1216
Oy	793	gccacgtcccaagagcacaagaatgcttcttcctctctatatggaatcatctgcaccat	852
Dd	1217	CCGCCATCCTCTGTGGAGCAGCAAGAAGTCTCTTTTCTATTATTTCACGTCTTTTCACCCCGAT	1276
Oy	853	gctgatcccttatatatataacttagaacacaaggatlaagaaagctttaaag	909
Dd	1277	GCTAAACCCTTGATTTTACAACCTTAGAAGTCTAGAGTCAAGGGTCCCTTAGAGAG	1333
RESULT	3		
A1604386/c			
LOCUS	A1604386	523 bp	mRNA linear EST 21-APR-1999
DEFINITION	vvt4c09.x1 Stratiogene mouse skin (#937313) Mus musculus cDNA clone IMAGE:1228144 3' similar to SW:OLF2_HUMAN Q15062 OLFACTORY RECEPTOR-LIKE PROTEIN FAT1L.; mRNA sequence.		
ACCESSION	A1604386		
VERSION	A1604386.1	GI:4613553	
KEYWORDS	EST.		
SOURCE	house mouse.		
ORGANISM	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
AUTHORS	1 (bases 1 to 523) Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person ,B., Swaller,T., Gibbons,M., Page,D., Harvey,N., Schurk,R., Ritter ,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and Wilson,R. The WashU-NCI Mouse EST Project 1999 Unpublished (1999)		
TITLE	Contact: Marra M/WashU-NCI Mouse EST Project 1999		
JOURNAL	Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA Tel: 314 286 1800 Fax: 314 286 1810		
COMMENT	Email: mouselst@wustl.edu This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information. MGI:653736 This clone was previously sequenced on the 5' end only, this new data is from the 3' end Possible reversed clone: similarity on wrong strand High quality sequence stop: 457. Location/Qualifiers 1..523 /organism="Mus musculus" /strain="C57BL/6" /db_xref="taxon:10090" /clone_image="IMAGE:1228144" /clone_lib="Stratiogene mouse skin (#937313)" /sex="females" /issue_type="whole skin" /dev_stage="11 weeks old" /lab_host="SODR (kanamycin resistant)" /note="Organ: skin; Vector: pBluescript SK-; Site:1; EcorRI ; Site:2; XhoI; cloned unidirectionally. Primer: Oligo dT. Whole skin from 11 week old C57BL/6 female mice. Average insert size: 1.0 Kb; Uni-ZAP XR Vector; ~5' adaptor sequence: 5' GAATTCGACGAG 3'-3' adaptor sequence: 5' CTCGACTGTCTTTTCTTTTTTTT 3'"		
BASE COUNT	149 a 133 c 150 g 89 t		2 others
ORIGIN			
Query Match	21.7%; Score 204.6;	DB 9;	Length 523;
Best Local Similarity	62.2%; Pred. No. 7.4e-45;		
Matches 321; Conservative	0; Mismatches 195;	Indels 0;	Gaps 0;
Oy	267	ggaagtattatcgctatogctgctgtgtgccccagctttatattctggccttgaggg	326
Dd	516	GAAAGACATCGCTTCTTGTGGATGTCACAGCTCTTCACTTATCTTGGGGA	457

OY	337		ctacgaatattcctccgagcgcaagtcttcgttatgttttgtaactatctgcgc	386
Db	456	CGAGGAGTGCATCTCCTGTACGACGATGGCCCTTTGACCCTATGTGGCTGTGGCCAGC		397
OY	387	ctctccattactcatgttalcatacgaccagaagactctgcctccagtltygaactgcattct	446	
Db	396	CCNMGCACTATGACCACCATTCATMCACCCTCGCCCTGTGTGGACAGCTGACACTGTGGCCT	337	
OY	447	gggttacatcggttttagtaattcaacgagtggtgtgtgtctaacctcacctctccagltgcacattc	506	
Db	336	GGACGATTGGCCTAAGTAGAGTGACGAGTTCAGAGGCCATTCACCTGTGGCTGCCTCTTCT	277	
OY	507	gtgacctccatctgtgatatacattctctgttgaagtcgccacactgctccaagtatctt	566	
Db	276	GTCGCCACACATCAGGTTBACGACCTTTGTGTGTGAAGTCCCTGCTTTGATGTGACTGTCT	217	
OY	567	gtgttgagacaacagcaaatgaggctgtaactatctctgtcagltgagtccttccatcaa	626	
Db	216	GTGGGAGACACCACTATTAATGATACAAATGGCATTTGCCAGTGTCTTCATCTTGATG	157	
OY	627	tacccttgaactatctcttatacataatgcttttatgttctgcgagcaatatlgagaaac	686	
Db	156	TACCTCTAGGCTCATCTTGTCTCTTATGGTCCCATTTGCCAGGGCACTGCTGAGAGATA	97	
OY	687	agtcctgcgaaggtctgcaaaaagcaatcttgtagacatggtctccatctaattgtggt	746	
Db	96	GCTCTGCAAAAGGGGGGCAAGAAGCTTTTGGAGCTGCTCTTCCACCTCATTTGGTCA	37	
OY	747	ctctttttatagtaagcgcgtctctgtgtaacctgc	782	
Db	36	CTCTCTCCACAGCTCAGTCACTCGCTGTATTATCTGC	1	
RESULT	4			
AZ090606/c				
LOCUS	AZ090606	692 bp	DNA	linear
DEFINITION	RPCI-23-27B20.TV RPCI-23 Mus musculus genomic clone RPCI-23-27B20,			
KEYWORDS	DNA sequence.			
ACCESSION	AZ090606			
VERSION	AZ090606.1	GI:7732649		
SOURCE	GSS.			
ORGANISM	Mus mouse.			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
AUTHORS	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murineae; Mus.			
JOURNAL	1 (bases 1 to 692)			
COMMENT	Zhao,S., Nieman,W., Feldblum,T., Malek,J., Shatsman,S., Akiret and Fraser,C.M. Mouse BAC End Sequences from Library RPCI-23 Unpublished (1999) Other_GSSs: RPCI-23-27B20.TJ Contact: Shaying Zhao Department of Eukaryotic Genomics The Institute for Genomic Research 9712 Medical Center Dr., Rockville, MD 20850, USA Tel: 301 838 0200 Fax: 301 838 0208 Email: szhaod@igf.org Clones are derived from the mouse BAC library RPCI-23. For BAC library availability, please contact Pieter de Jong [pietere@jmg.med.buffalo.edu]. Clones may be purchased from BACPAC Resources (http://bacpac.med.buffalo.edu/orderingframe.htm) or from Resea ch Genetics (info@resgen.com). BAC end page: http://www.tlgr.org/tldb/bac_ends/mouse/BAC_end_intro.html Plate: 27 row: B column: 20 Seq primer: T7 Class: BAC ends. Location/Qualifiers 1..692 /organism="Mus musculus" /strain="C57BL/6J"			

```

/db_xref="taxon:10090"
/clone="RPCI-23-27B20"
/clone_lib="RPCI-23"
/sex="Female"
/lab_host="DH10B"
/note="Organ: Kidney/Brain; Vector: pBAC3.6; Site:1;
ECORI; Site:2; EcoRI; Female C57BL/6J mouse kidney and/or
brain genomic DNA was isolated and partially digested
with a combination of EcoRI and EcoRI Methylase. Size
selected DNA was cloned into the pBAC3.6 vector at the
EcoRI sites. The ligation products were transformed into
DH10B electrocompetent cells (BRL Life Technologies)."
BASE COUNT      205 a      164 c      163 g      160 t
ORIGIN

```

```

Query Match      21.2%; Score 200.2; DB 12; Length 692;
Best Local Similarity 58.0%; Pred. No.1.3e-43;
Matches 374; Conservative 0; Mismatches 268; Indels 3; Gaps 1;

```

```

Oy 279 gttatcgtgctgtagccagcttccatcttctgagcctgggagctactgaatc 338
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 692 GCATGCTGGGTGATGACCCAGTCTTCTATACACTCTTACGGGGAAGAGTGTG 633

Oy 339 ttctcctgagcgtcatgctccttgatgttgtagctatttgaggcctccattact 398
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 632 TGCTCCTGTGTGATGGCTTTCGACCGGTATGCTGTGTGCTGCTGCTGCTGCTG 573

Oy 399 cagttatcagcagagagactcgtcctccagctgagcagctgagctgagctgagct 458
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 572 CCAGCATTTATGACCCCTTCTCTGCCATGATGAGCCATCTCTCTGAGTGGAGGCC 513

Oy 459 ttagtaactagtggtgtgtgtctacacactcctcagctcagctcagctcagctc 518
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 512 TTGTGAACCTCTGACTGACAGCAAGTCTCATATGACATCACTCTGTGGCCATCACC 453

Oy 519 tgaataatactctctctgtgaagtcctcagcagctcagctcagctcagctcagct 578
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 452 TGA--ACCACTTCTCTGTGAGATGCTTGTCTCTGGAAGCTGTGAGGACACAC 396

Oy 579 cagcaaatgagcgtgagactcctcgtcagtgagctcctcctcctcctcctcctc 638
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 395 TGGGAACAGAGACCTTCTTGTGGCTGGAGCTGTATTTGTGCTGTGCTGTAGCAC 336

Oy 639 tcaatcctatcatatgcttcttattgtccgagcagctatgagatagcagctcgtga 698
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 335 TAATTTAGGACACCTTATGCAACATTTGCTATGCTGATGAGATCAAGTCAAGATCT 276

Oy 699 gtccacaataaagattgtgagacatgtgtgtccatctaatgtgtgtgtgtgtgt 758
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 275 GCGCAGAAAGGCTGTGGAACTTGTGGGTCCCACTTACTGTGTGTGTGTGTGTGT 216

Oy 759 gtacacgctcctcgtgtacacacacacacacacacacacacacacacacacac 818
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 215 GCTCAGCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 156

Oy 819 tgggttctctctctctctctctctctctctctctctctctctctctctctct 878
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 155 TTGCTCCTCCTCTTTTACTATATCTATCTATCTATCTATCTATCTATCTATCT 96

Oy 879 ggaacagagagagagagagagagagagagagagagagagagagagagagag 923
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 95 GAACAAGAGATGTAAAGGGGCTCTGTGCAAGGCTACTTGTAAAG 51

```

```

RESULT 5
A2969227
LOCUS A2969227 642 bp DNA linear GSS 27-APR-2001
DEFINITION 2M04J1J24R Mouse 10kb plasmid UUGC2M library Mus musculus genomic
ACCESSION A2969227
VERSION A2969227.1 GI:13840454
KEYWORDS GSS.

```

```

SOURCE house mouse.
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 642)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beecorn,T., Duval,B., Hamli,C.,
Islam,H., Longacre,S., Mahmoud,M., Meene,E., Pedersen,T., Reilly
M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausen,A.
and Wright,D., Weiss,R.
TITLE Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
JOURNAL Unpublished (2000)
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLG, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0241 row: J column: 24
Seq primer: CACACAGAAACAGCATGACAC
Class: plasmid ends
High quality sequence stop: 642.
Location/Qualifiers
1..642
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC2M024J1J24"
/clone_lib="Mouse 10kb plasmid UUGC2M library"
/sex="Female"
/lab_host="E. coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PMD42uv; Purified genomic DNA from M.
musculus C57BL/6J (female) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adaptor DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of PMD42 (q11473211419b1AF129072.1), a copy-number
inducible derivative of plasmid RL. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adaptor mouse DNA was annealed to
adaptor vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

```

```

BASE COUNT      130 a      166 c      129 g      217 t
ORIGIN

```

```

Query Match      20.9%; Score 197.2; DB 12; Length 642;
Best Local Similarity 56.7%; Pred. No.8.1e-43;
Matches 364; Conservative 0; Mismatches 278; Indels 0; Gaps 0;

```

```

Oy 234 tccacaatgctagtaattatgcagcatcaggaagaatcagctatcgtgtgtgtg 293
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1 TCCCCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 60

Oy 294 tagccagcttccatctcgtcgtgagcctcgtcgtcgtcgtcgtcgtcgtcgtc 353
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 61 TGACCCAGCTCTTTGGCATTCATTTTCTTTGTTGGCTCAGAGTCTCTCTCTGCGACAA 120

Oy 354 tctccttgattgttgtagctattgttcgagcctccattcattcagctatcagctcacc 413
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 121 TGGCATATGATGATATATTTGATCTGTAAGCCGTTAAGGATCTGATTTATATGACAA 180

```


QY	414	agagactctgcctccagcttggcgaagtccatgctggtaacggttttagtaagcaccaggt	473
Db	181	AGGCTCTGTGACGCTGGTTACGACGGTTTCATCTGGAACAGGTGGGGTTTCTCAACTAGTGT	240
QY	474	ggtgtctaccctctactctccagctccagctccagctctgagcccatctatgtatagatcattcc	533
Db	241	TGCACACGGCTTTTGACCTTTCACCTGACCTGCCCTTTTGGTAAACATCAGATTAATTTCT	300
QY	534	tctgtgaagctccctgcaactgtctcaagtatctgtgtgtgagacaacagcaatgagcgctg	593
Db	301	TCTGTGACATCACTCCCTTGGCTCATCTGTCTGTGGGGATGACTTCCCTCATGAACTGG	360
QY	594	aacattctctgtcagtgtagctctctccatctctaaataccccctgaactatccttatatat	653
Db	361	CTTTGCTGTCCATTGGAGTCCCTCATAGGCTGTGACCTCTTCTCTGTGCTCATCTCTTCT	420
QY	654	atgctcttatgtctcgaagcagtatgtgaagatacagctctgcctgaagctgtgacaanaagat	713
Db	421	ACCTTTACATCATCTCCACACATCCTCTGAGGATCCGTTCTCTGAGGGAGGACACAAAGCT	480
QY	714	ttagagacatgtgtgtcccatctaatgtgtgtgtctcttttataagtaacagcgctctctg	773
Db	481	TTTCCACCTGTGGCTCCACACCTGCTCATTTGTTATCTCTATTATGACAGTGCTATCTTCA	540
QY	774	tgtactctgaaccacacactctgcgccagctcccaaggacaaggaagaatgtctctctctc	833
Db	541	CGTATGTGAGGCCCATCTCATCTTACTCTCAGAGAAAAGATGATGTGATCTCAGTGTGT	600
QY	834	atggaatcatctgcaaccacatgctgtaatccccctatatatacag	875
Db	601	ATATGTTGTTCACACACCATGCTGAATCTCTGTAATTTATTAACGC	642
RESULT	6		
LOCUS	AK017036	1394 bp	mRNA linear
DEFINITION	Mus musculus adult male testis cDNA, RIKEN full-length enriched library, clone:993433E02:similar to T1 OLFACCTORY RECEPTOR, full insert sequence.		
ACCESSION	AK017036		
VERSION	AK017036.1	GI:12856091	
KEYWORDS	HTC; CAP trapper.		
SOURCE	Mus musculus (Strain:C57BL/6J) adult male testis cDNA to mRNA, clone_11b:RIKEN full-length enriched mouse cDNA library clone:493343E02.		
ORGANISM	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
AUTHORS	Carninci, P. and Hayashizaki, Y.		
JOURNAL	High-efficiency full-length cDNA cloning		
MEDLINE	Meth. Enzymol. 303, 19-44 (1999)		
PUBMED	99279253		
REFERENCE	10349636		
AUTHORS	2 (sites)		
TITLE	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.		
JOURNAL	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes		
MEDLINE	Genome Res. 10 (10), 1617-1630 (2000)		
PUBMED	20495374		
REFERENCE	11042159		
AUTHORS	3 (sites)		
TITLE	Shibata, K., Itoh, M., Aizawa, K., Nagoka, S., Sasaki, N., Carninci, P., Kono, H., Akiyama, J., Nishi, K., Kitsumai, T., Tashiro, H., Itoh, M., Suni, N., Ishii, T., Nakamura, S., Hazama, M., Nishino, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiike, S., Inoue, K., Togawa, Y., Iwawa, M., Ohara, E., Watabiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.		
JOURNAL	RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer		
MEDLINE	Genome Res. 10 (11), 1757-1771 (2000)		
PUBMED			
REFERENCE			
AUTHORS			

	MEDLINE PUBMED REFERENCE	10576913 11076861 4 (sites)
	AUTHORS	The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
TITLE	JOURNAL	Functional annotation of a full-length mouse cDNA collection Nature 409, 685-690 (2001)
REFERENCE	AUTHORS	5 (bases 1 to 1394) Aachi,T., Aizawa,K., Akahira,S., Akimura,T., Aono,H., Arai,A., Arakawa,T., Baldarelli,R., Bono,H., Brownstein,M., Bult,C., Carninci,P., Fukuda,S., Fukushima,Y., Furuno,M., Hanagaki,T., Hara,A., Hayatsu,N., Hill,D., Hiramoto,K., Hirooka,T., Horii,F., Humé,D., Imclant,K., Ishii,Y., Itoh,M., Izawa,M., Kasukawa,T., Kato,H., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Okazaki,Y., Okido,T., Owa,C., Quackenbush,J., Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H., Saeki,D., Schriml,L., Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T., Soejabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,I., Tanaka,T., Tejima,Y., Toyota,Y., Yamamura,T., Yamakoshi,T., Yasunishi,A., Yoshida,K., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.
TITLE	JOURNAL	Direct Submission Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physiscal and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp, url:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)
COMMENT		Please visit our web site (http://genome.gsc.riken.go.jp/) for further details. cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science laboratory in RIKEN Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was plimed with a primer [5' GAGAGAAGAGATCCACAGACGTTCCTTTTCTTTTTTAA 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 10 and subtraction to Rot = 100.0. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAAGATTCCTCAGATTAAATTAATTAATCCCCCCCCCCC 3']. cDNA was cleaved with BamHI and XhoI. cDNA of size comprised between 0.5 and 3 kb was selected before cloning. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC I. Cloning sites, 5' end: SalI; 3' end: BamHI. Host: DH10B. Location/Qualifiers
FEATURES	SOURCE	1. .1394 /organism="Mus musculus" /strain="C57BL/6J" /db_xref="MGD:MGI:1894709" /db_xref="taxon:10090" /clone="4933433EO2" /sex="male" /tissue_type="testis" /clone_lib="RIKEN full-length enriched mouse cDNA library" /dev_stage="adult" 332..1282 /note="data source:SPRR, source key:09EPF8, evidence:ISS putative similar to J1 OLFACTORY RECEPTOR" /codon_start=1 /protein_id="BAB30564.1" /db_xref="GI:12856092" /translation="MEPWNTLESGLFIILDGSGPELLCAVTLLYMLALISNGI, LLIVITADARLHVPMYLLRLSLIDLIPLFSVYPNVVDVLADNLTISEGCALOQIF, SAMTGGADEDLIAFNAYDRRYVAICHLNPIWFPSPACRAMVAISWLTSIGRTYS, VYMHPPECSOEIRHLCIEVPRLKACADTGYELAMVVTVGGIIPILPSLAITSY, SLLPFLYLMRPSNDEKKALKLVTCSSHTLVVGMTFGCATPMVLPSPSHSKODNIISV, FYIVYPALNPFLYLSENKVYAIGAVRVRDGRHIIIPAHATV"
CDS		355 a 337 c 304 g 398 t
PAGE COUNT		

ORIGIN				
Query Match 20.5%; Score 193.4; DB 11; Length 1394;				
Best Local Similarity 52.3%; Pred. No. 1,1e-41;				
Matches 428; Conservative 0; Mismatches 391; Indels 0; Gaps 0;				
OY	104	tacacgtgacacatctctggaatctgacatctatcttagtgcagccctggaacacaa	163	
DB	434	TACATGTGGACATGCATGACATGAGACTTCTACTCTGTCATCAGACAGTGGATCCGG	493	
OY	164	cttaacaccccatgattttctctaccatctatctaccccgatcttgtaacc	223	
DB	494	CTTACGATACCCATGATACCTCTCTACTGAGCAGATGTCCTCATTTGACCTCTCTCACA	553	
OY	224	acacgtacagtcacacaaatgctagtaattatgacagcatcagaagaatcatgatat	283	
DB	554	TCAATTTGAACTCCCAACACAGTGTGATTTTCTGCTCAGACAGAACACATATCTTT	613	
OY	284	cgtggtcgtgtagccagcttctcatattctggccttgagggtactgaatctctc	343	
DB	614	GAGGAGATGCCCTTCAATTTCTCAGCAATGACATTTGGTGGTGACAGAGACCTCTCT	673	
OY	344	ctggccctcatgctcttctgattggttgtagctatttggtgacctctccatctcagtt	403	
DB	674	CTGGCCTTCAATGACCTTATGATGATGATGAGTGTGCTCATTTGCTCAATCTTAACTCATGATC	733	
OY	404	atcatgacacagagacatctgctccagttgagcagctgcacatctggttactgattagt	463	
DB	734	TTCTATGATCCAAAGGCTGACAGGCTCATGATGATGATGATGATGATGATGATGATGAT	793	
OY	464	aactcagtggtgtgtctaccctgactctccagctgcacatctgtagccatctgtgata	523	
DB	794	AGTCCCTTAGTGCACACAGTGTACACATGACATGCTCCCTTTTCCATGATGCCAGAAATC	853	
OY	524	gatacctctctctgtagagctcctgacgtgctcaagtattcttgtaggaacagca	583	
DB	854	AGACACCTGCTCTGTGAGTCTCTCCATTTGTAATTTGGTTGTGACAGACATCTTCAA	913	
OY	584	aatgagcgtgacatctctctgtagagctctctcaataataacccctgacatctc	643	
DB	914	TATGAGCTCATGTTTATGATGACAGGATGATTTCTTATGCTCCCTCTGCTGCATTT	973	
OY	644	cttatcatatgctttattatgtccgagcagatgtaggaatacagctgcctggaagtcga	703	
DB	974	ATTACCTCTACTACTCAATTAATTTATTCACATGCTGACATGCTTCAAAATGAGGCGAG	1033	
OY	704	caaaaacatttggagacatggtgtcccatctaatgtgtgtctcttttatagtaca	763	
DB	1034	AAGAAAGCCCTGTGACCTGTCTCTCCACATGATGATGATGATGATGATGATGATGATG	1093	
OY	764	gcgcgtctctgtagcctgcaacacacctgcgcagctccaaagcaagaagaatggt	823	
DB	1094	GCCACTTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	1153	
OY	824	tctctctcatatgtaataatgacacatgctgaatcccttatataatacctagaac	883	
DB	1154	TCTGTGTTTATATACATTTGTCACACAGCTCTGAACCCCTCATTTACAGCTGAGAAAT	1213	
OY	884	aaggaagtaagaagaagcttaaaaggttggttcaaga	922	
DB	1214	AAGGAGGTGATGAGCTGTGTAAGAGAGTACTGGGAGA	1252	

RESULT 7
 AK016560
 LOCUS 3063 bp mRNA linear HTC 19-JAN-2002
 DEFINITION Mus musculus adult male testis cDNA, RIKEN full-length enriched library, clone:4932441H21; similar to T1 OLFATORY RECEPTOR, full insert sequence.
 ACCESSION AK016560
 VERSION AK016560.1 GI:12855357
 KEYWORDS HTC; CAP trapper.

SOURCE	Mus musculus (strain:C57BL/6J) adult male testis cDNA to mRNA, clone:4932441H21.
ORGANISM	Mus musculus
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS	1 (sites)
TITLE	Carninci,P. and Hayashizaki,Y.
JOURNAL	High-efficiency full-length cDNA cloning
MEDLINE	99279253
PUBMED	10349636
REFERENCE	2 (sites)
AUTHORS	Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Kono,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE	20499374
PUBMED	11042159
REFERENCE	3 (sites)
AUTHORS	Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Kono,H., Akiyama,Y., Nishi,K., Kitamura,T., Tashiro,H., Itoh,M., Suni,N., Ishii,Y., Nakamura,S., Hazama,M., Nishino,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kasahara,K., Yonijawa,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watanabe,M., Fieda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuda,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kita,A. and Hayashizaki,Y.
TITLE	RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multipipillary sequencer
JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE	20530913
PUBMED	11076861
REFERENCE	4 (sites)
AUTHORS	The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium
TITLE	Functional annotation of a full-length mouse cDNA collection
JOURNAL	Nature 409, 685-690 (2001)
MEDLINE	5 (bases 1 to 3063)
PUBMED	Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Aono,H., Arai,A., Arkawa,T., Baldarelli,R., Bonio,H., Brownstein,M., Bult,C., Carninci,P., Fukuda,S., Fukunishi,Y., Furuno,M., Hanagaki,T., Hara,A., Hayatsu,N., Hill,D., Hiramoto,K., Hiraoka,T., Hori,F., Hume,D., Imocani,K., Ishii,Y., Itoh,M., Izawa,M., Kasukawa,T., Kato,H., Kawai,J., Kojima,Y., Kono,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Nishi,K., Nomura,K., Nunakata,R., Ohno,M., Okazaki,Y., Okido,T., Owa,C., Quackenbush,J., Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D., Schirral,L., Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Tanaka,Y., Tejima,Y., Toyai,T., Yamamura,T., Yamahaka,I., Yasunishi,A., Yoshida,K., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.
TITLE	Direct Submission
JOURNAL	Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp, URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)
COMMENT	Please visit our web site (http://genome.gsc.riken.go.jp/) for further details.
	cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5' GAGAGAGACGACGACCAAGACGCTCTTTTCTTTTCTTTTNN 3']. cDNA was prepared by using cDNAase thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 10.0 and subtraction to Rot = 100.0. Second strand cDNA was prepared with the primer

TITLE , Durick, K. and Pollok, B.
JOURNAL Unpublished tags from a HL-60 Genomescreen(TM) Library
COMMENT Unpublished (2000)
Contact: Greg Henkel
Gene Expression
Aurora Biosciences Corp.
11010 Torreyana Road, San Diego, CA 92121, USA
Tel: 8584048436
Fax: 8584046719
Email: henkel@aurorabio.com

Pools of cells were isolated from a Genomescreen(TM) library. The library of cells was generated by retroviral integration of a gene tagging element consisting of: 1) A promoterless beta-lactamase gene preceded by a splice acceptor as a reporter for gene expression; 2) A promoter driving neomycin resistance followed by a splice donor to trap downstream exons. 3' RACE from neomycin gene was performed using total RNA from isolated pools. Output was shotgun cloned in pAMP-1 and used to transform DH5-alpha competent bacteria. 5' ends of reported sequences were immediately preceded by splice donor from the trapping construct.

Class: exon-trapped.
Location/Qualifiers
1. .326
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="Genetrap HL-60 Human Promyelocytic Leukemia Library"
/tissue_type="acute promyelocytic leukemia"
/cell_type="promyeloblast"
/note="Organ: peripheral blood; Vector: pAMP-1; 3' RACE of total RNA from genetrap pools; shotgun clone in pAMP-1 and used to transform DH5-alpha competent bacteria."

BASE COUNT 109 a 65 c 57 g 95 t

ORIGIN

Query Match 19.5%; Score 184; DB 12; Length 326;
Best Local Similarity 100.0%; Pred. No. 2.5e-39;
Matches 184; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 760 taagacgcgtctctgttctgacgtgcaacacacctgcccagctcccaaggaagcaagaagat 819
|||||
Db 1 TACAGCGCTCTCTGTACTGCAACACACTCTCCCAAGCTCCCAAGCAAGAAAGAT 60
|||||

QY 820 gtttctctctctctctgaatcaltgacccatgctgaatcccttataatacacttag 879
|||||
Db 61 GGTTCCTCTCTATGGAATCATTCGACCCCATGCTGAATCCCTTATATATACACTTAG 120
|||||

QY 880 gaacaaggaagtaagaagagcttaaaaggttggtgcaagagctctcttaacaagaa 939
|||||
Db 121 GAACAAAGAGGTAAAGAGAGCTTTAAAGGTTGTCAGAGACTCTTTAAATCAAGAA 180
|||||

QY 940 ataa 943
|||||

Db 181 ATAA 184

RESULT 11
AZ516219 640 bp DNA linear GSS 16-OCT-2000
LOCUS AZ516219/c
DEFINITION RPCI-11-350G23, TVK RPCI-11 Homo sapiens genomic clone
ACCESSION RPCI-11-350G23, DNA sequence.
VERSION RPCI-11-350G23, GI:10824405
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 640)
Zhaio, S., Adams, M.D., Nierman, W., Malek, J., de Jong, P. and Venter, J.C.

TITLE BAC end sequences of library RPCI-11
JOURNAL Unpublished (1997)
COMMENT Contact: Shanying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@tigr.org

Clones are derived from the human BAC library RPCI-11. For BAC library availability, please contact Pieter de Jong (pieter@delong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from Research Genet cs (info@resgen.com). BAC end search page: http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html. This BAC end was generated during the R&D process and may have higher chance of clone tracking errors.

Seq primer: T7
Class: BAC ends.
Location/Qualifiers
1. .640
/organism="Homo sapiens"
/db_xref="GDB:7634182"
/db_xref="taxon:9606"
/clone="RPCI-11-350G23"
/clone_lib="RPCI-11"
/sex="Male"
/cell_type="Lymphocytes"
/note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI; RPCI11 Human Male BAC library"

BASE COUNT 185 a 99 c 140 g 216 t

ORIGIN

Query Match 18.9%; Score 178.6; DB 12; Length 640;
Best Local Similarity 93.0%; Pred. No. 9.2e-38;
Matches 198; Conservative 0; Mismatches 14; Indels 1; Gaps 1;

QY 732 atctaattggtgtgtctcttttatagtacagcgtctctgttacctgaaccacct 791
|||||
Db 640 ATCTAATGTGTGTGTCATCTTTTATGTACAGCATCTCATGTACCTGCACACACTT 581
|||||

QY 792 cggccagctccaaggaagcaagaagatggtttctctcttataaggaatattgaccca 851
|||||
Db 580 CACCACGCTCCAAAGACCGGGGAAAGATGTTCTCTCTTGTGGAAATTCACCCCA 521
|||||

QY 852 tgcgaatcccttatataatacacttaagaaaggaagtaagaagcttaaaagt 911
|||||
Db 520 TGCATGATCCCTTATATATACACTTAGAACAAGAGTAAGAAACCTTTAAAGCT 461
|||||

QY 912 tggttgc-aagagctctcttaatacaagaataa 943
|||||
Db 460 TGGTTGCAAGAGTCTTCTTATATCAAGAAATTA 428
|||||

RESULT 12
BH11304 788 bp DNA linear GSS 19-JUL-2001
LOCUS BH11304
DEFINITION RPCI-24-367N6, TU RPCI-24 Mus musculus genomic clone RPCI-24-367N6,
DNA sequence.
ACCESSION BH11304
VERSION BH11304, GI:14946319
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Soturognath; Muridae; Murinae; Mus.
1 (bases 1 to 788)
Zhaio, S., Nierman, W., Malek, J., Shatsman, S., Akiret, B., Levins, M., Tsegaye, G., Geer, K., Krol, M., Shvartsbeyn, A., Gebregeorgis, E., Russell, D., de Jong, P. and Fraser, C.M.
Mouse BAC End Sequences from Library RPCI-24
Unpublished (1999)

COMMENT Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@tigr.org
Clones are derived from the mouse BAC library RPCI-24. For BAC library availability, please contact Pieter de Jong (pdejong@mail.cho.org). Clones may be purchased from BACPAC Resources (<http://www.choi.org/bacpac/orderingframe.htm>). BAC end plate: http://www.tigr.org/tldb/bac_ends/mouse/bac_end_intro.html
Seq primer: SP6
Class: BAC ends.

FEATURES
Location/Qualifiers
1..788
/organism="Mus musculus"
/strain="C57Bl/6J"
/db_xref="taxon:10090"
/clone="RPCI-24-367N6"
/clone_1lb="RPCI-24"
/sex="Male"
/cell_type="Spleen/Brain"
/note="Vector: pFARBAC1; Site_1: BamHI; Site_2: BamHI; RPCI-24 Mouse BAC Library produced by Pieter de Jong. The library was cloned in the pFARBAC1 cloning vector at the BamHI sites using MboI partially digested male C57Bl/6J DNA."

BASE COUNT 168 a 217 c 162 g 241 t
ORIGIN

Query Match 18.8%; Score 177.2; DB 12; Length 788;
Best local Similarity 53.7%; Pred. No. 2.4e-37;
Matches 411; Conservative 0; Mismatches 353; Indels 2; Gaps 2;

QY 104 tacactgtgacacatcttggcaatctgacatattctagtgtaacgctggagaccacaa 163
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 6 TACATCTTCAGCCTGCGCAATGGAATGTAATCTAGTACATCTGCGACCCCAAG 65

QY 164 cttcataccccctgtatttttcttaccatctacactcctcgtgattctgttaacc 223
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 66 CTGCACACCCCTTATGATCTTCTTTCATCTGCGCATCTCTGACATATCCTATGCT 125

QY 224 acatgtacagtcaccacaatgctagtaattatgacagcagcaagaaagtaacagttat 283
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 126 TCCAAACAATGTCCCAAGATGTTGTCCAACCTTATAACCAAGAAAGAACCAATCTATT 185

QY 284 cgtgctgtgttagccagcttctcataattctgacctgggggactgaataatctctc 343
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 186 GCCCATGATCATACACAGACATCTTGATTTGGCCCTTGCCTTGGAGTGGCTGAT 245

QY 344 ctggcgcgtatgtccttggattgtgt-actatttgcgcgccttcataactcagt 402
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 246 TTGGCAGCCATGTCATATGACAGTTTGGGCACTCTGCCACCCCTTACATACACTGT 305

QY 403 tatcatgacacagagactcgtccctcagttggcagctcactccttggttactgttttag 462
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 306 CATCATGAGTTGGAAGATATGTGTGGCCCTGGCTGTCTCATCTTGGATGTGATTTAG 365

QY 463 taactcagtggtgtgtctaacctactctcagctgagctcactctgtgacccctatgtat 522
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 366 CCTCTGTTGGACACACATCTCTCTGTAAGGTTGCCCTTTTGTGGGCCCGAGAAAT 425

QY 523 agtactccttctgtgaaagtcctcgtgacgtcctcaagttatcttctgtgttagaagaagc 582
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 426 AAACCACTCTTCTGTGAATCTGTGGCTGTCTCAAGCTGGCCCGCAGACACTTTGAT 485

QY 583 aataggagctgaactatctcttgcagtgagccttcacatataaccctgagcaactca- 641
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 486 CAACCAATGTCATCTGACTGATGTGTTTTCGTTAGTTGGACCCCTTTGCTTACG 545

QY 642 tccatatacatatgcttttattgtccgagcagtaattgagatcacgtctgtaaggtc 701
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 546 TGCTTGTATCTTACACCCACATCCTCTGACAAATCTGAAATGCAATCAAGAGAGGTC 605

QY 702 gacaaaagcatcttggagacatggttcccatctcaattggtgtgtctttttatagta 761
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 606 GCAGGAAGCCCTTTCCACCTGTCCTCCATCTCTGTGTGGCTCTTCTTTGGTA 665

QY 762 cagcgtctctgtgtacactgcaaccaccttcgcagctcccaagacaaagaaagatg 821
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 666 TAGCATTGCTTGTGTTTATGCTCCCTGACTGATGATCAGAGAGAGAGAGAAAT 725

QY 822 ttctctcttctatgataatcagaccacatgctgaatccccctat 867
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 726 TTGTCTACGTGTCACAGTCTTTTACCAGAGGATTCCTCAT 771

RESULT 13
BH331857
LOCUS BH331857 853 bp DNA linear GSS 03-DEC-2001
DEFINITION CH230-46N15, TV CHORI-230 Segment 1 Rattus norvegicus genomic clone
CH230-46N15, DNA sequence.
ACCESSION BH331857
VERSION BH331857.1 GI:17262571
KEYWORDS GSS.
SOURCE Norway rat.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 853)
Zhao,S., Shetty,J., Shatsman,S., Tsengaye,G., Geer,K., Shvartsbeyn
A., Gebregeorgis,E., Overton,L., Russell,D., Chen,D., Riggs,F., de
Jong,P. and Fraser,C.M.
Rat BAC End Sequences from Library CHORI-230 EcoRI segment
Unpublished (1999)
Other_GSSs: CH230-46N15.TJ
Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@tigr.org
Clones are derived from the rat BAC library CHORI-230
(<http://www.choi.org/bacpac/rat230.htm>). For BAC library
availability, please contact Pieter de Jong (pdejong@mail.cho.org).
Clones may be purchased from BACPAC Resources
(http://www.choi.org/bacpac/ordering_information.htm). BAC end
plate: http://www.tigr.org/tldb/bac_ends/rat/bac_end_intro.html
Seq primer: T7
Class: BAC ends.

FEATURES
Location/Qualifiers
1..853
/organism="Rattus norvegicus"
/strain="BN/SSNHsd/MCw"
/db_xref="taxon:10116"
/clone="CH230-46N15"
/clone_1lb="CHORI-230 Segment 1"
/sex="Female"
/cell_type="Brain"
/note="Vector: pFARBAC2.1; Site_1: EcoRI; Site_2: EcoRI;
CHORI-230 Rat (BN/SSNHsd/MCW) BAC library produced by
Pieter de Jong"

BASE COUNT 210 a 172 c 158 g 313 t
ORIGIN

Query Match 18.7%; Score 176; DB 12; Length 853;
Best local Similarity 52.4%; Pred. No. 5.1e-37;
Matches 409; Conservative 0; Mismatches 370; Indels 1; Gaps 1;

OY	108	cttggacacatccttttggcaaatctgacaatatattctagtgtaacgccttggaaaccaacttc	167
Db	61	CTGATACAGGTGGTGGGAACAATTCTCAATAAATCGGGACTGCTTTTATTCCAACCCTCC	120
OY	168	atacccccatgtatttttcttccaaactatcatactcctctgacctgttgttaaccacat	227
Db	121	ATFACACCTATGTACTCTTCTTCATTTCCAACTGTCTTTTATTGACATCTGTTTTTACACCA	180
OY	228	gtacagtcgcccaaaatgctagtaaattatgcaagcatcaaggaaaaglaactagtagcg	287
Db	181	CCACCATACCAAATAATGTTGGTGAAATATTACAGACTCAGATCAGATCAGCTCCATFTAGTTATGTAG	240
OY	288	gctgtgtgcccgaagctttcatattctcgtgccttgggggctactgaaatacttctccrg	347
Db	241	GCTGCCTCACCAAAATCTGCTTTGTTTGTAGCTTTTGTGACGAAATGGAAATTCGTGG	300
OY	348	cagtatcttcccttgatttgatttgagtattttgacggcctccatcacatcagaftatca	407
Db	301	TAAATATGGCTTAATGATATAGTTTGCTGATCTGTCATCTGAGTACACTGTGCATTA	360
OY	408	tgcacagaagaactctgcctcagatttgcagctgcacgcctcctgggttactgtttttagaact	467
Db	361	TGAATCCCAAAATTCGTGTGTGATGAGCTTTGTGTCTCTTCTTGATTAAGCATTTCTAGATG	420
OY	468	cagtgatgtgtctaaccttgaactctcgaactcgaactgtgcaacttctgtgacccotatgtatagatc	527
Db	421	CTTTGCTCCACACTTTGATGGCAGCTGCCCTGTCATTTCTGCACAAAGGTGAAATTCCTCG	480
OY	528	aattctcttggaaagtcctctgcacgcctcgaagtatcttctgtgtgagcaacagaagt	587
Db	481	ACTTTTGTGGAACTAGCTCATATTCTCAAGCTTGCAAGTGTCCAATTTCTAATCAATA	540
OY	588	aggcgtgaactatctccttgtagtgccttccatcctaatacccttgacaactcatcetta	647
Db	541	AATATCTGGTGATTTGGTGACAAAGTCTGTTGGGTCTCTTCCACTGTCTGTATTAATTT	600
OY	648	tatcatatgcttttatctgbcgagcagatgatgagatacagtcctgtcgaaggtcgacaaa	707
Db	601	AATCTTAATCTAGAATTTATTTCTCTGTGTTGAAAATTCATCCAGCTCAGAAAAATATA	660
OY	708	aagcatcttggagcagatgtgttccactccaattgtgtgtctctttttataglaagcgg	767
Db	661	AGGTTTCTCCACATGTGCTTACACTTAGTAGTAGTGATCTTGTTCTAATGGTACAGGTT	720
OY	768	tctctgtgtaccttgaaccacacttcgcccagctcgaagaccaaagaaagaatgyttctc	827
Db	721	TTGGGTATTATGTAGGTTCTGCTGGCACTCATAGACACTAGAAAGAGTGCAATAGCCTCTG	780
OY	828	tcttcta-tgnaatcatgtgacccaatgctgtaatcccttatatalacactltagaacaag	886
Db	781	TGATGTACACTGTGTGCTACCTCCATGATGAATCTTTTATATATAGCTTTGAGAAATAAG	840
RESULT 14			
LOCUS	AZ413817	588 bp	DNA linear GSS 03-OC7-2000
DEFINITION	JM0188F05F Mouse 10kb plasmid UUC1M library Mus musculus genomic clone UUC1M0188F05 F, DNA sequence.		
ACCESSION	AZ413817		
VERSION	AZ413817.1	GI:10537830	
KEYWORDS	GSS.		
SOURCE	Mus musculus house mouse.		
ORGANISM	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 588)		
AUTHORS	Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamll,C., Islam,H., Longacre,S., Mahmood,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D., Weiss,R.		
TITLE	Mouse whole genome scaffolding with paired end reads from 10xb plasmid inserts		
JOURNAL	unpublished (2000)		

COMMENT	CONTACT: ROBERT B. WEISS	UNIVERSITY OF UTAH	BIOMEDICAL POLYMERS RESEARCH BLDG., 20 S. 2030 E., SLC, UT 84112, USA	TEL: 801 585 5606	FAX: 801 585 7177	EMAIL: ddm@genetics.utah.edu	INSERT LENGTH: 10000	STD ERROR: 0.00	PLATE: 0188	ROW: F	COLUMN: 05	SEQ PRIMER: CGTGTGTAACAGCAGCGCACG	CLASS: PLASMID ENDS	HIGH QUALITY SEQUENCE STOP: 588.
FEATURES	Location/Qualifiers													
SOURCE	1..588													
	/organism="Mus musculus"													
	/strain="C57BL/6J"													
	/db_xref="taxon:10090"													
	/clone="UUCGIM0188F05"													
	/clone_lib="Mouse 10kb plasmid UUCGIM library"													
	/sex="Male"													
	/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"													
	note="Vector: PMD42HV; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource													
	(http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of PMD42 (911473211419b1AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."													
BASE COUNT	157	a	139	c	104	g	188	t						
ORIGIN														
Query Match	18.6%; Score 175.6; DB 12; Length 588;													
Best Local Similarity	66.2%; Pred. No. 5.8e-37;													
Matches 253; Conservative	0; Mismatches 129; Indels 0; Gaps 0;													
QY	32	gagttattctgctggtgttctcagaatgacctgsgcgagttcccaactctctgtgtc	91											
DB	198	GACCTTCATCTCTCTGGGATTCTTCACGTGACCCCAACGAGACACATCATCTCTGCACTT	257											
QY	92	tctctgatcttcttaccactgtgacacactctcttgcaactcgcacattatctagtgacgc	151											
DB	256	GCTCTTGCTGCTTGTGATATGTGTGACTGTGGTAGGAACACACACATATTCTAGTGTCCAT	317											
QY	152	ctgsgacacacaaactcacaaccoccatgtatatttttcttaccacatcatcaactctcgat	211											
DB	318	CTGGACCTCCACACTCCATATCTCCATGTATTTCTCTTATCTAAATTTGCTTTGCTGGAC	377											
QY	212	ctttgttaccacactgtgacagttcccaacaatgtcagtaaatattatgcaagatcaagaa	271											
DB	378	CTCTTTTATACACTAGCAGCTTGTCTCCACAGATGCGTGAATATCTATGGGACCCCAAGAG	437											
QY	272	gtaacacgttaccgtgctgtgtgacacagcttcttcatattcttgcctgtgggctact	331											
DB	438	TCTATTACCTACGAGGAGGGGTGTACTCCAGTTCTTCTTTGCCCTGAGACTGGAGCACCA	497											
QY	332	gaatatctctcccgccgcatgctcctttgattgtgtgttagctattgttcgcgctctc	391											
DB	498	GAATGCTCTCTCTTGCGCTGATGGCCATAGACCCGCTATGGCTGTGCTCAACCTCTT	557											

THIS PAGE BLANK (USPTO)